PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES
Volume 22 May 15, 1936 Number 5

# *A SECOND LAW OF THE MOTIONS OF ERUPTIVE PROMINENCES*

By Edison Pettit

Mount Wilson Observatory, Carnegie Institution of Washington

Communicated April 2, 1936

Prominences which rise from the chromosphere and disappear at great heights above the sun during intervals of a few minutes to a few hours have been called eruptive. The nature of this more-or-less vertical motion was first determined from observations[1] of the great prominence of May 29, 1919, and a study of all the earlier data[2] and the observations obtained since[3] has abundantly confirmed the principle that eruptive prominences rise with uniform motion, modified at intervals by sudden increases as though occasional impulses had been given to them. This characteristic is illustrated by figure 1, where No. 1 is the time-height diagram for the prominence of May 29, 1919, and No. 38 is a similar plot for the prominence of Oct. 21, 1934.

We now have such diagrams for thirty-eight eruptive prominences for which four or more observations have been made, scattered over the interval since Trouvelot made the first measures visually in 1885. Seven of the plots show no breaks at all, while as many as four breaks are found in the remainder.

Figure 1 illustrates a simple relationship between the increasing velocities of eruptive prominences which appears to be quite general for the entire list. It will be noted that for No. 1 the first velocity, 5.5 km./sec., is almost tripled by the impulse indicated by the first break in the plot; that the new velocity of 14.7 km./sec. is almost doubled by the next impulse; and this new velocity of 27.9 km./sec. is almost exactly doubled by the last impulse, the final observed velocity being 60 km./sec. For No. 38 the first velocity of 13.5 km./sec. is almost exactly tripled (observed value, 40 km./sec.) by the single impulse indicated by the break in the curve. The observed velocity ratios are not exactly whole numbers, but they are surprisingly close approximations, in view of the nature of the material. If we assume that in No. 1 the initial velocity was 5 km./sec., thus implying an observational error of 0.5 km./sec., the succeeding velocities would be 15,

TABLE 1

VELOCITIES $V_0$ AND ACCELERATIONS $A_c$ IN ERUPTIVE PROMINENCES

| No. | $V_0$ | $O-C$ | $F$ | $Q$ | $A_c$ | $F_a$ |
|---|---|---|---|---|---|---|
| 1 | 5.5 | +0.6 | 1 | *F* | 4.93 | 1 |
| | 14.7 | −0.1 | 3 | *E* | 9.9 | 2 |
| | 27.9 | −1.7 | 2 | *G* | 14.8 | 3 |
| | 60.0 | +0.8 | 2 | *F* | 29.6 | 6 |
| 2 | 37.0 | −3.7 | 1 | *F* | 40.74 | 1 |
| | 163.9 | +0.9 | 4 | *E* | 122.1 | 3 |
| 3 | 4.0 | −1.2 | 1 | *F* | 5.15 | 1 |
| | 26.0 | +0.2 | 5* | *G* | 20.6 | 4 |
| 4 | 11 | +0.6 | 1 | *F* | 10.32 | 1 |
| | 19 | −0.6 | 2 | *G* | 10.3 | 1 |
| | 42 | +0.7 | 2 | *P* | 20.6 | 2 |
| 5 | 7 | −0.2 | 1 | *F* | 7.20 | 1 |
| | 22 | +0.4 | 3 | *G* | 14.4 | 2 |
| | 43 | −0.2 | 2 | *G* | 21.6 | 3 |
| 6 | 6 | −0.2 | 1 | *F* | 6.19 | 1 |
| | 31 | 0.0 | 5* | *F* | 24.8 | 4 |
| | 54 | .... | .. | *NG* | .... | .. |
| 7 | 7 | +0.4 | 1 | *NG* | 6.60 | 1 |
| | 37 | +4.0 | 5* | *F* | 26.4 | 4 |
| | 64 | −2.0 | 2 | *F* | 33.0 | 5 |
| 8 | 3 | +0.1 | 1 | *F* | 2.90 | 1 |
| | 29 | 0.0 | 10* | *G* | 26.1 | 9 |
| 9 | 64 | +0.3 | 1 | *F* | 63.7 | 1 |
| | 191 | −0.1 | 3 | *G* | 127.4 | 2 |
| | 317 | .... | .. | *NG* | .... | .. |
| 10 | 14 | +0.5 | 1 | *NG* | 13.5 | 1 |
| | 54 | 0.0 | 4 | *F* | 40.5 | 3 |
| 11 | ... | .... | .. | ... | .... | .. |
| 12 | 53 | +6.8 | 1 | *P* | 46.2 | 1 |
| | 89 | −3.4 | 2 | *G* | 46.2 | 1 |
| 13 | 2.2 | .... | .. | *NG* | .... | .. |
| | 83.3 | −3.4 | 1 | *F* | 86.86 | 1 |
| | 175.5 | +1.8 | 2 | *P* | 86.9 | 1 |
| 14 | 70 | −3.9 | 1 | *G* | 73.9 | 1 |
| | 223 | +1.6 | 3 | *G* | 147.8 | 2 |
| 15 | 82 | +6.9 | 1 | *NG* | 75.1 | 1 |
| | 223 | −2.3 | 3 | *F* | 150.2 | 2 |
| 16 | 400 | .... | 1 | *G* | 400 | 1 |
| 17 | 119 | −5.4 | 1 | *F* | 124.4 | 1 |
| | 375 | +1.8 | 3 | *G* | 248.8 | 2 |
| 18 | 15 | +0.6 | 1 | *NG* | 14.4 | 1 |
| | 144 | 0.0 | 10* | *E* | 129.6 | 9 |
| 19 | 400 | .... | .. | *G* | 400 | 1 |
| 20 | 357 | .... | .. | *G* | 357 | 1 |
| 21 | 23 | .... | .. | *P* | 23 | 1 |
| 22 | 4.4 | +0.8 | 1 | *F* | 3.58 | 1 |
| | 6.6 | −0.6 | 2 | *F* | 3.6 | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 28.3 | −0.3 | 4 | *P* | 21.5 | 6 |
| | 57.2 | −0.1 | 2 | *F* | 28.6 | 8 |
| | 115.0 | +0.4 | 2 | *P* | 57.3 | 16 |
| 23 | 8 | 0.0 | 1 | *F* | 8 | 1 |
| | 136 | 0.0 | 17* | *G* | 128 | 16 |
| | 59 | .... | .. | *G* | .... | .. |
| 24 | 71 | .... | .. | *P* | 71 | 1 |
| 25 | 2 | +0.06 | 1 | *G* | 1.94 | 1 |
| | 20 | +0.6 | 10* | *G* | 17.5 | 9 |
| | 155 | −0.1 | 8 | *G* | 135.8 | 70 |
| | 13 | −0.6 | 1 | *G* | 13.65 | 1 |
| | 82 | +0.1 | 6 | *F* | 68.2 | 5 |
| 26 | ... | .... | .. | *NG* | .... | .. |
| 27 | ... | .... | .. | *NG* | .... | .. |
| 28 | 81 top | .... | .. | *G* | .... | .. |
| | 200 top | .... | .. | *G* | .... | .. |
| | 103 bottom | .... | .. | *G* | .... | .. |
| | 176 bottom | .... | .. | *G* | .... | .. |
| | 92 av. | −3.2 | 1 | *G* | 95.2 | 1 |
| | 192 av. | +1.6 | 2 | *G* | 95.2 | 1 |
| 29 | 4 | .... | 1 | *P* | 4 | 1 |
| 30 | 45 | +6.0 | 1 | *P* | 39.0 | 1 |
| | 75 | −3.0 | 2 | *P* | 39.0 | 1 |
| 31 | 3 | −0.1 | 1 | *G* | 3.1 | 1 |
| | 19 | +0.4 | 6* | *G* | 15.5 | 5 |
| | 37 | −0.2 | 2 | *G* | 18.6 | 6 |
| 32 | 21 | .... | .. | *F* | 21 | 1 |
| 33 | 86 | .... | .. | *F* | 86 | 1 |
| 34 | 4 | −0.8 | 1 | *G* | 4.16 | 1 |
| | 25 | 0.0 | 6* | *G* | 20.8 | 5 |
| 35 | 5 | +0.4 | 1 | *G* | 4.63 | 1 |
| | 19 | +0.5 | 4* | *E* | 13.9 | 3 |
| | 74 | −0.1 | 4 | *G* | 55.6 | 12 |
| | 126 | .... | .. | *E* | .... | .. |
| | 105 | .... | .. | *G* | .... | .. |
| 36 | 3.9 | +0.5 | 1 | *G* | 3.42 | 1 |
| | 15.5 | +1.8 | 4* | *G* | 10.3 | 3 |
| | 54.2 | −0.5 | 4 | *P* | 41.0 | 12 |
| 37 | 155 | .... | .. | *G* | 155 | 1 |
| | 75 | −2.0 | 1 | *F* | 77.0 | 1 |
| | 155 | +1.0 | 2 | *G* | 77 | 1 |
| 38 | 13.5 | +0.2 | 1 | *G* | 13.35 | 1 |
| | 40 | 0.0 | 3 | *G* | 26.7 | 2 |

30 and 60 km./sec., with errors $(O-C)$ of +0.5, −0.3, −2.1 and 0.0 km./sec., respectively, and the corresponding accelerations at the breaks in the curves are 5, 10, 15 and 30 km./sec., respectively. A simple least-squares solution shows that if we assume an initial velocity of 4.93 km./sec. the errors $(O-C)$ are still further reduced (table 1). These sequences of values

suggest that when the velocity of an eruptive prominence changes, its new velocity is the old one multiplied by a small whole number and that the acceleration is therefore also a small multiple of the initial velocity.

To test the generality of this principle, all the velocities for the thirty-eight cases are analyzed in table 1. The first two columns give the catalog identity of the prominence, as published in the references, and the observed velocity $V_0$. The value of the initial velocity $V$ for each case was computed by a simple least-squares solution; the next velocity was then found by applying to $V$ the factor $F$ in the fourth column opposite that velocity,

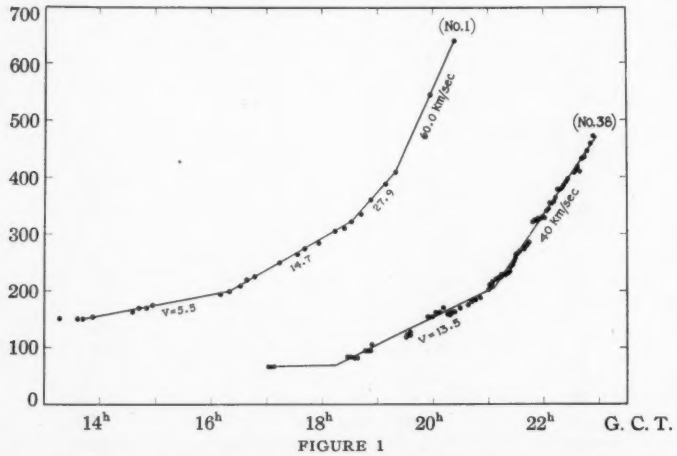



FIGURE 1

Motions of eruptive prominences of May 29, 1919 (No. 1) and Nov. 14, 1934 (No. 38); heights in thousands of kilometers, ordinates in time (G. C. T.). Note that the ratios of successive increasing velocities $V$, with good approximation, are small whole numbers.

and so on. By this means the residuals $(O-C)$ were found for the various velocities of the prominence. The quality of the observed velocity $V_0$, derived from an inspection of the plots and given in the column $Q$, is rated as excellent, $E$; good, $G$; fair, $F$; poor, $P$; and not dependable, $NG$. The initial acceleration is of course $V_0$, or more probably $V$ if we suppose that there is no smaller velocity; the succeeding accelerations, computed from the changes in computed velocity, are found in the sixth column. These values are all multiples of the initial velocity $V$, the successive factors $F_a$ being given in the last column.

The column of residuals $(O-C)$ shows that, in general, the observed velocities are represented within the errors of measurement by a law which states that *when the velocity of an eruptive prominence changes, the new velocity is a*

*small multiple of that immediately preceding it.* It should be noted that the law as stated is much more exacting than the requirement that each velocity in a sequence be merely a multiple of the initial value.

Numbers 1, 22 and 25, in which several changes in velocity occur, are quite convincing. Two outstanding exceptions are found in No. 28 and in the last two velocities for No. 35. For No. 28 both the top and the bottom of the rising detached mass were measured. One change in velocity was found in both plots, that for the top somewhat more than doubling in velocity and that for the bottom falling considerably short of doubling. If, however, we take the mean, which applies to the center of gravity of the moving mass (opposite the symbol "av." in the table) the ratio is almost exactly 2. This probably means that where a volume of gas is rising its changes in volume are superposed upon the velocity of the prominence, which has a different origin but the same property of uniform motion. In this case the prominence first contracted slightly, then, upon changing velocity, began to expand. No. 1 is a similar case, but the measures were made originally on the center of gravity. In No. 35 rapid expansion of the head of the prominence began at the break in the curve representing the next to the last velocity. It is possible that the increase in velocity of the crest observed at this point is due to this expansion. The last velocity is that of a small knot which remained after the head had faded and may not represent its average motion.

In table 1 the values of $V_0$ are for the most part tabulated to the nearest kilometer as they appear on the plots given in the references. The use of decimal fractions in the tables does not materially change the result. Spacing in the columns of Nos. 25, 28 and 37 indicates that two points were measured, yielding two independent sets of data. An asterisk in the $F$ column indicates that this factor is so large and the initial velocity so small that an error of 1 km./sec. in the initial velocity would change the factor by one unit. Therefore little confidence can be had in this value of $F$. Probably some of the very low initial velocities are largely the ordinary growth of the prominence. There seems to be little indication of a fundamental initial velocity common to all eruptions; if any, 5 km./sec. seems to be the most frequent factor.

[1] *Astrophys. Jour.*, **50**, 206 (1919).
[2] *Pub. Yerkes Observatory*, **3**, Part IV (1925).
[3] *Mt. Wilson Contr.*, No. 451; *Astrophys. Jour.*, **76**, 9 (1932).

# THE EFFECT OF AUXIN ON THE ABSCISSION OF PETIOLES[1]

BY CARL D. LA RUE

DEPARTMENT OF BOTANY, UNIVERSITY OF MICHIGAN

Communicated April 7, 1936

Several years ago the writer noticed a tree of *Betula alba* var. *papyrifera* (Marsh.) Spach., on which almost every leaf had been attacked by caterpillars. These larvae had eaten away almost every part of the blades of the leaves, only the mid-rib and the main veins remaining, but the leaves did not fall. Instead, the wounds were healed with a periderm layer, and these skeletonized leaves lived at least the greater part of the season. Afterward, similar leaves were observed on *Alnus incana* (L.) Moench., where, also, they did not abscis as might have been expected, but persisted through the summer. Such behavior is in striking contrast to that of leaves, from which the entire blades have been removed leaving only the petioles which, in most plants, absciss within a few days.

The hypothesis, which had at first occurred to the writer, that the wound caused by the removal of the leaf lamina might in some way stimulate the development of the abscission layer failed to take account of the behavior of the skeletonized leaves of Betula and Alnus. Furthermore, experiments in which wounds approximately 100 times as extensive as those caused by severing the upper ends of the petioles were made in leaves of several species of trees and shrubs proved that even such severe injury would not cause abscission.

Disturbed water relations due to the sudden cessation of transpiration were considered as a possible cause of petiole abscission but were shown to be inadequate by the discovery that the retention of even a small portion of the leaf blade was enough to prevent the fall of the mutilated leaf. Also, experiments were performed in which normal and mutilated leaves were confined under bell jars in an atmosphere saturated with water vapor. If petioles deprived of their blades became ingested with water because of root pressure or water flow from any other cause, and this condition was able to cause stimulation of the abscission layer, similar conditions should be set up in petioles with a small amount of leaf blade left intact and kept in a saturated atmosphere to reduce transpiration. However, it was found that the reduced transpiration within the saturated atmosphere under the bell jars did not cause abscission in *Coleus blumei*, either of entire leaves, or of those in which the leaf blades had been reduced to $^1/_2$, $^1/_4$, $^1/_8$, $^1/_{20}$ or $^1/_{40}$ of their normal size. (The fact, previously observed by the author,[3] that leaves of *Populus grandidentata* Michx., *P. tremuloides* Michx. and other species confined in a saturated, unventilated atmosphere

soon fall off, needs explanation but does not invalidate the results of the foregoing experiments.)

It did not appear that the cause of the fall of bladeless petioles could be due to reduced photosynthetic activity, because the small amount of leaf lamina which sufficied to retain the petioles on the stem could scarcely be significant in food production.

The cause of the abscission of the bladeless petioles seemed to be the lack of some substance produced in the leaf blades which was effective even in very small quantities such as could be produced by minute fragments of leaf blade. The observed facts seemed to indicate that, contrary to the first ideas of the writer, abscission always resulted when the development of the abscission layer was not inhibited by some substance produced in the leaf blades, and that the abscission layer did not need any special stimulus for its development.

TABLE 1

ABSCISSION OF PAIRS OF PETIOLES OF COLEUS IN WHICH BOTH PETIOLES HAVE RECEIVED THE SAME TREATMENT

| NUMBER OF EXPERIMENT | NUMBER OF PAIRS OF PETIOLES | TREATMENT GIVEN PETIOLES | AVERAGE NO. HOURS BEFORE ABSCISSION "A" PETIOLES | AVERAGE NO. HOURS BEFORE ABSCISSION "B" PETIOLES |
|---|---|---|---|---|
| 1 | 6 | Blocks of plain agar applied to cut ends | 68.0 | 68.0 |
| 2 | 6 | Blocks of agar containing ether extract of Populus pollen applied to cut ends | 52.5 | 50.0 |
| 3 | 6 | Blocks of agar containing exudate of bases of Vallota leaves applied to cut ends | 65.7 | 65.7 |
| 4 | 4 | Blocks of agar containing pollen of Populus applied to cut ends | 87.0 | 111.0 |
| 5 | 2 | Blocks of agar containing pollen of Populus applied to cut ends | 120.0 | 120.0 |

*Trials with Auxin-Containing Substances.*—It was already known that the auxin produced by the stem tip inhibits the growth of axillary buds, as has been shown by Snow[4] and Thimann and Skoog.[5] It was known also from the work of Thimann,[6] and Thimann and Went,[7] that auxin is formed in the leaves and flows downward through the petiole to the stem. These facts led the writer to the assumption that auxin was the substance most likely to serve as the inhibitor of the abscission layer. This assumption was tested by a series of experiments which are presented in table 2. Only after all but three of these experiments were completed, did the author discover the note by Laibach[2] of the effect of the growth-hormone in orchid pollen on petiole abscission in Coleus.

In these experiments opposite pairs of petioles of *Coleus blumei* Benth.

TABLE 2

THE EFFECT OF AUXIN-CONTAINING SUBSTANCES ON ABSCISSION OF PETIOLES OF COLEUS

| NO. OF EXPERIMENT | NO. OF PAIRS OF LEAVES | EXPERIMENTAL TREATMENT | TREATMENT OF CONTROLS | AV. NO. OF HOURS BEFORE FALL OF EXPERIMENTAL PETIOLES | AV. NO. OF HOURS BEFORE FALL OF CONTROL PETIOLES | % OF DIFFERENCE |
|---|---|---|---|---|---|---|
| 1 | 4 | Agar blocks containing Coleus leaf exudate | Blocks of plain agar | 58.5 | 43.5 | + 34.5 |
| 2 | 4 | Agar blocks with exudate of leaves of Vallota purpurea | Blocks of plain agar | 123.0 | 75.0 | + 64.0 |
| 3 | 6 | Lanolin mixed with Populus grandidentata pollen | Untreated lanolin | 63.0 | 45.0 | + 40.0 |
| 4 | 6 | Agar blocks with exudate of Vallota purpurea | Blocks of plain agar | 52.3 | 35.7 | + 46.5 |
| 5 | 6 | Lanolin with pollen of Populus grandidentata | Untreated lanolin | 120.0 | 58.0 | +106.9 |
| 6 | 6 | Lanolin with pollen of Populus grandidentata | Untreated lanolin | 188.0 | 131.3 | + 43.2 |
| 7 | 4 | Agar blocks with exudate of Vallota purpurea | Blocks of plain agar | 56.0 | 39.0 | + 43.6 |
| 8 | 9 | Lanolin with exudate of Vallota purpurea | Untreated lanolin | 89.3 | 38.8 | +129.7 |
| 9 | 11 | Agar mixed with pollen of Populus grandidentata | Blocks of plain agar | 80.7 | 70.9 | + 14.0 |
| 10 | 10 | Agar mixed with pollen of Populus grandidentata | Blocks of plain agar | 87.0 | 102.6 | − 15.6 |
| 11 | 3 | Lanolin mixed with spores of Equisetum arvense | Untreated lanolin | 61.0 | 39.0 | + 56.5 |
| 12 | 5 | Lanolin mixed with pollen of Podocarpus macrophylla | Untreated lanolin | 43.6 | 43.6 | 0.00 |
| 13 | 6 | Lanolin with 2% urine | Untreated lanolin | 61.0 | 41.0 | + 48.8 |
| 14 | 5 | Lanolin with 2% urine | Untreated lanolin | 84.0 | 58.0 | + 44.8 |
| 15 | 7 | Lanolin with 2% urine | Untreated lanolin | 70.8 | 38.8 | + 82.4 |

were used; one of each pair, "*A*," being used for experimental treatment and the other, "*B*," as control. The essential uniformity of abscission rate of pairs of petioles when both receive the same treatment is shown in table 1. This appears also in Experiments 9, 10 and 12 in table 2 where the experimental treatments were ineffective.

With the exception of three experiments, the trials recorded in table 2 show that auxin-containing substances, when applied to the cut end of the petiole, have the ability to delay abscission to a significant extent. Experiment 12 seems to show that Podocarpus pollen does not contain auxin in sufficient quantity to inhibit abscission. The pollen of *Populus grandidentata* when mixed with lanolin has inhibiting properties which seem to be absent when it is mixed with agar. It is possible that the mixing of the pollen with melted agar produces the equivalent of a water extract which has been shown by Thimann[6] to allow a rapid oxidation of the auxin. However, a study of the whole set of experiments shows: that only Experiments 9 and 10 were performed in full light in the greenhouse; that Experiment 12 was carried out in the shade of a greenhouse bench; and that all the other experiments were made in the dark room. Apparently, light may play a significant part in the process. Further mention will be made of this subject in another part of this paper.

*Experiments with Synthetic Hetero-Auxin.*—Although the results given in table 2 indicate a strong probability that it is the auxin contained in the various substances which inhibits the abscission of the petioles, they give no certainty that some other substance may not be responsible. As a means of securing more definite limitation of the abscission effect the writer secured, through the kindness of Professors F. W. Went and K. V. Thimann of the California Institute of Technology, a supply of the pure synthetic hetero-auxin. This was mixed with lanolin, 0.5 mg. hetero-auxin to 100 cc. melted lanolin. The lanolin was then forced into rubber tubes of 5-mm. bore and 2.5 cm. in length. For controls, untreated lanolin was used, also in rubber tubes. When the leaf blades were cut off, the rubber tubes were slipped over the upper ends of the petioles. In this way it was possible to hold an adequate amount of lanolin on the petiole even if the room temperature became rather high. The results of treatment of petioles with pure lanolin are given in table 3.

From Experiment *A* as shown in table 3, it is evident that hetero-auxin is able to inhibit the development of the abscission layer even more effectively than any of the mixtures used in the experiments given in table 2. Even though only one experiment has been performed, the results are so striking that probably we are safe in concluding that the substance which prevented petiole abscission in the experiments given in table 2, and in that of Laibach, was auxin.

The striking lack of effect of the hetero-auxin in Experiments *B* and *C* of

TABLE 3

THE EFFECT OF HETERO-AUXIN ON PETIOLE ABSCISSION IN COLEUS

| EXPERIMENT | NO. OF PAIRS OF PETIOLES | TREATMENT OF EXPERIMENTAL PETIOLES | TREATMENT OF CONTROL PETIOLES | NO. OF HOURS BEFORE ABSCISSION OCCURRED | | % OF DIFFERENCE |
|---|---|---|---|---|---|---|
| | | | | EXPERIMENTAL PETIOLES | CONTROL PETIOLES | |
| A | 15 | Hetero-auxin in lanolin (in dark room) | Untreated lanolin (in dark room) | 130.70 | 54.00 | +141.98 |
| B | 39 | Hetero-auxin in lanolin (in full light in greenhouse) | Untreated lanolin (in full light in greenhouse) | 182.51 | 159.76 | + 14.24 |
| C | 30 | Hetero-auxin in lanolin (in full light in greenhouse) | Untreated lanolin (in full light in greenhouse) | 105.80 | 103.86 | + 1.88 |

table 3 remains to be explained. In some way the presence of light prevents the action of the hetero-auxin. It may be that in strong light the petioles themselves may form enough auxin to prevent abscission for as long a time as the experimentally applied auxin. Some support is given this idea by the rather long persistence of the control petioles in Experiments *B* and *C*, which is greater than that observed in any other control petioles in this series of experiments.

Another possibility is suggested by the results of Van Overbeek as related by F. A. F. C. Went.[8] Van Overbeek found a decreased sensitivity toward growth substance of the cells of seedlings of Raphanus which were exposed to light. Perhaps in a similar way the effect of hetero-auxin on the cells of the abscission layer is diminished by light, though here we would have a decreased inhibition instead of a decreased elongation. No attack has yet been made on the question as to why auxin should inhibit the development of the abscission layer, or the mechanism of the action.

*Trials with Ricinus Communis.*—An experiment with 20 pairs of petioles of Ricinus was made in which the experimental petioles were treated with 2 per cent urine in lanolin and the controls with pure lanolin. The experiment was run in a dark room and showed an increase of experimental petioles over the controls of 35.2 per cent. In a second experiment involving 15 pairs of petioles, the experimental petioles were treated with 0.5 mg. hetero-auxin in 100 cc. of lanolin and pure lanolin was applied to the control petioles. This experiment was made in deep shade un-

der a greenhouse bench, and showed an advantage of 36.12 per cent of the experimental over the control petioles. Because the leaves are alternate in Ricinus, pairs of petioles of identical age cannot be secured. Possibly this accounts for the small effect secured by treatment of the petioles with urine and hetero-auxin. The disadvantage can be overcome by the use of large numbers of petioles, and Ricinus, like Coleus, is more favorable for use than many other plants because petioles deprived of their laminae are abscissed within a few days. While the advantages gained by auxin and hetero-auxin treatment are smaller in these experiments than in most of those in Coleus they are consistent with them and serve at least to show that the effect of the growth hormone in inhibiting petiole abscission is not confined to a single species of plant.

*Summary.*—Petioles of Coleus from which the blades have been removed absciss within a few days. Several substances, known to contain the growth hormone, applied in agar or lanolin to the severed ends of the petioles, are able to inhibit the development of the abscission layer for significant lengths of time. Synthetic hetero-auxin in lanolin, applied to the cut ends of the petioles, inhibits abscission in Coleus for a longer time than any of the auxin-containing substances.

In strong light, neither the auxin-containing substances nor the pure hetero-auxin are able to delay abscission.

Results with petioles of Ricinus are in accord with those performed on Coleus.

[1] Papers from the Department of Botany, University of Michigan, No. 563.

[2] Laibach, F., *Ber. deut. Bot. Ges.*, **51**, 386–392 (1933).

[3] La Rue, C. D., *Amer. Jour. Bot.*, **20**, 1–17 and 159–175 (1933).

[4] Snow, R., *Proc. Roy. Soc. London*, **B108**, 305–316 (1931).

[5] Thimann, K. V., and Skoog, F., *Ibid.*, **B114**, 317–339 (1934).

[6] Thimann, K. V., *Jour. Gen. Physiol.*, **18**, 23–24 (1934).

[7] Thimann, K. V., and Went, F. W., *Proc. Konink. Akad. Wet. Amsterdam*, **37**, 456–467 (1934).

[8] Went, F. A. F. C., *Biol. Reviews*, **10**, 187–207 (1935).

# PERISSODACTYLA OF THE SESPE EOCENE, CALIFORNIA

BY CHESTER STOCK

BALCH GRADUATE SCHOOL OF THE GEOLOGICAL SCIENCES, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated April 14, 1936

*Introduction.*—One of the outstanding peculiarities of the Eocene faunas known from the Sespe deposits of California is the continued absence of any evidence of the Equidae. Aside from lack of representatives of this family, however, the Perissodactyla include members of the Rhinocerotoidea and Brontotheriidea. Two types have already been recorded.[1] It is the purpose of the present paper to draw attention to additional fossil material belonging to the Tapiroidea and Rhinocerotoidea.

## Family Helaletidae: Dilophodon, sp.

*Material.*—A single upper molar, either $M\underline{2}$ or $M\underline{1}$, No. 1949, Plate 1, figure 1.

*Locality.*—Tapo Ranch, Locality 180 C.I.T. Vert. Pale. Sespe Upper Eocene.

*Description.*—This tooth furnishes as yet the only evidence of presence of the family Helaletidae. The specimen is smaller than comparable teeth in *Helaletes boöps* and *H. nanus* and comes closest in size to *Dilophodon minusculus* (No. 1634a A.M.N.H.) from the Bridger Middle Eocene of Wyoming. It represents an animal apparently not so small as *Heteraletes leotanus* Peterson from the Duchesne River horizon of Utah.

Comparing the Sespe specimen with a molar tooth of *D. minusculus*, No. 1634a A.M.N.H., the following characters may be noted. A greater disparity prevails between the levels of the tip of the paracone and the tip of the metacone. The depth of the cross-lophs, as viewed from the inner side, appears to be slightly lower. The cingulum at the outer base of the metacone is not so strongly developed as in *D. minusculus*. A cingulum is present on the anterior and on the posterior sides of the tooth. Measurements (in millimeters) of specimen No. 1949 are: anteroposterior diameter 8.6; greatest transverse diameter normal to inner side 9.2.

## Family Hyracodontidae: Triplopus? woodi, n. sp.

*Type Specimen.*—No. 476 C.I.T. Vert. Pale. Coll., Plate 1, figures 3, 3*a*, a fragmentary ramus with $D\overline{p3}$, $D\overline{p4}$, $M\overline{1}$ and $M\overline{2}$. The species is named for Dr. Horace Elmer Wood, 2nd, in recognition of his valuable suggestions in the course of study of the Sespe Perissodactyla.

*Referred Specimen.*—Fragment of ramus with $M\overline{3}$, No. 652.

*Locality.*—Brea Canyon section of the Sespe, north of the Simi Valley, Ventura County, California, Locality 128 C.I.T., approximately 120 feet lower stratigraphically than Locality 150. The referred specimen, No. 652, occurred at Locality 148 C.I.T., approximately 10 feet lower stratigraphically than Locality 150.

*Specific Characters.*—Similar in characters of jaw and dentition to, but showing a progressive increase in size beyond, the Bridger and Lower Uinta species of *Triplopus*. Teeth more slender than in type of *Epitriplopus uintensis* (Peterson).

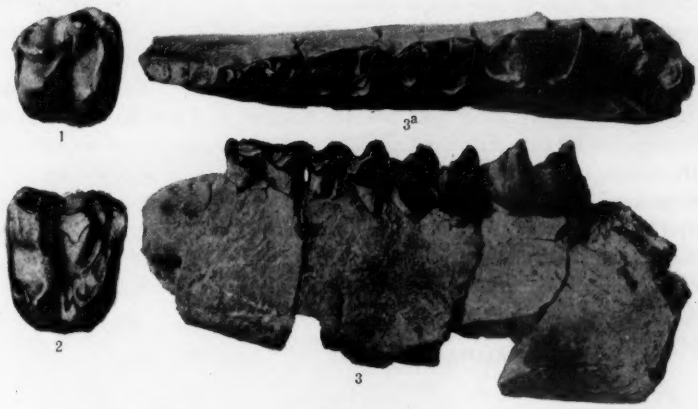



PLATE 1

*Dilophodon*, sp.

Figure 1, $M\underline{1}$ or $M\underline{2}$, No. 1949, occlusal view; × 2.

Triplopodine or Hyrachyid Rhinoceros

Figure 2, $D\underline{p4}$, No. 1950, occlusal view; × 1.

*Triplopus? woodi*, n. sp.

Figures 3, 3*a*, type specimen, No. 476, ramus with $D\underline{p3}$–$\underline{M2}$, inclusive; lateral and occlusal views; × 1.

Calif. Inst. Tech. Coll. Sespe Eocene, California.

*Description.*—No. 476, as indicated by the measurements, shows an appreciable increase in size beyond the *Triplopus* represented by the lower jaw, No. 2342 A.M.N.H. from the middle Eocene Bridger of the Washakie Basin, Wyoming. Considerable similarity prevails between these two in the dentition as well as in the proportions of the jaw. A posterior mental foramen is situated below the tooth socket of $D\underline{p2}$ in the Sespe jaw and has therefore a position slightly anterior to that in No. 2342. In both specimens this foramen is situated much closer to the ventral border than to the alveolar border.

The type of *T. cubitalis* Cope represents a young animal and in this specimen the teeth are distinctly smaller and the jaw shallower than in No. 476 from the Sespe (see comparative measurements).

In the referred specimen of *T. ? woodi*, No. 652, $M\overline{3}$ measures in anteroposterior diameter 17.4 mm. and in greatest transverse diameter approximately 10.3 mm. The depth of the jaw behind $M\overline{3}$ is 34.5 mm. This specimen is of interest because it extends upward the stratigraphic range of the triplopodine group in the Sespe deposits to within a few feet of Locality 150.

Increase in size of the Sespe individuals is likewise to be noted when comparisons are made between specimen 476 and 652 and the holotype and referred specimens of *Triplopus grangeri* Wood.[2]

In the lower jaw of the type of *Epitriplopus uintensis* the deciduous and permanent teeth are noticeably heavier than in *T. ? woodi*. Thus in the molars the width in relation to length of tooth is greater in the former than in the latter.

Relationship of the Sespe form to the triplopodine rhinoceroses is suggested but reference to one of the known genera from the American Eocene is not clearly established. The Sespe material is referred tentatively to *Triplopus*, although additional remains might show closer approach to *Epitriplopus*.

COMPARATIVE MEASUREMENTS (IN MILLIMETERS)

| | *Triplopus cubitalis* COPE TYPE NO. 5095 A.M.N.H. | *Triplopus* SP. NO. 2342 A. M. N. H. BRIDGER MID. EOCENE | *Triplopus? woodi* N. SP. TYPE NO. 476 C.I.T. SESPE |
|---|---|---|---|
| Length of tooth row $Dp\overline{3}$–$M\overline{2}$ | 40.4 | 52.8 | 56.7 |
| $Dp\overline{3}$, anteroposterior diameter | 9.3 | 12.1 | 12.4 |
| $Dp\overline{3}$, greatest transverse diameter | 5 | 6.5 | 6.8 |
| $Dp\overline{4}$, anteroposterior diameter | 8.7 | 11.5 | 11.7 |
| $Dp\overline{4}$, greatest transverse diameter | *a*6.3 | 7.6 | 8 |
| $M\overline{1}$, anteroposterior diameter | 9.1 | 12.8 | 14.1 |
| $M\overline{1}$, greatest transverse diameter | 6.7 | 8.2 | 9.3 |
| $M\overline{2}$, anteroposterior diameter | 11.3 | 15.6 | 16.6 |
| $M\overline{2}$, greatest transverse diameter | 7.2 | 9.3 | 10.2 |
| Depth of jaw at anterior end of $Dp\overline{3}$ | 14.9 | 18.7 | 24.4 |
| Depth of jaw at posterior end of $M\overline{2}$ | 16.9 | 23.8 | 30.2 |

*a* Approximate.

## Triplopodine or Hyrachyid Rhinoceros

A well-worn $Dp\underline{4}$, No. 1950, Plate 1, figure 2, represents a rhinoceros of uncertain affinities from Locality 150.

Measurements (in millimeters) of No. 1950 are as follows: Anteroposterior diameter 18.8; greatest transverse diameter normal to inner side 22.2. The tooth approaches in size the comparable deciduous molar

in *Hyrachyus? douglassi*, No. 1929 A.M.N.H., but differs in several noteworthy characters. In the Sespe specimen the outer and inner borders are roughly parallel, whereas in No. 1929, as also in No. 12137 A.M.N.H. of *Hyrachyus eximius*, the transverse diameter of the tooth is distinctly greater in front than in back. Moreover, the parastyle in these Bridger teeth is better developed and reaches farther externally with reference to the paracone than in the Sespe specimen. The external cingulum, present at the external base of the metacone in *Hyrachyus*, is not present in the Sespe specimen.

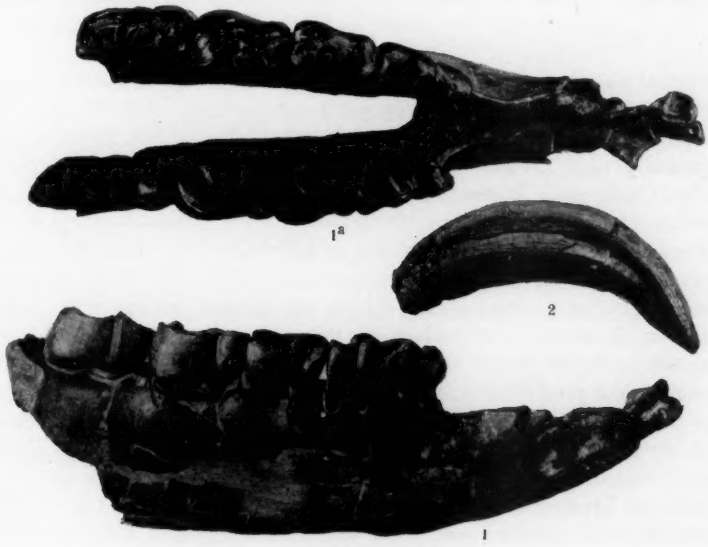



PLATE 2

*Amynodontopsis bodei* Stock

Figures 1, 1*a*, mandible, No. 2000, lateral and superior views; × 1/2. Figure 2, lower canine, No. 2001, lateral view; × 1/2.
Calif. Inst. Tech. Coll. Sespe Eocene, California.

Comparison of No. 1950 with the figure[3] of $Dp\underline{4}$ in the type of *Epitriplopus uintensis* shows that the tooth from the Sespe is of a different shape, being considerably wider for its anteroposterior diameter. Moreover, the parastyle is differently constructed in the latter and the internal cingulum on the anterior lobe is not so well developed as in *E. uintensis*.

$Dp\underline{4}$ in *Triplopus*, as represented by this tooth in No. 1973 A.M.N.H. resembles in shape No. 1950 from the Sespe. Aside from size, the principal difference rests in the fact that the parastyle reaches slightly far-

ther externally than in No. 1950, the outer border of the convexity being on a level with that of the paracone, whereas in the latter the outer wall of the convex parastyle lies slightly to the inner side of the outer wall of the paracone.

It is unfortunate that more material of this rhinocerotoid form is not available from Locality 150. Present evidence suggests perhaps a closer relationship to a triplopodine rather than to a hyrachyid type. In this connection it is interesting to recognize that the triplopodine rhinoceroses are known to range upward in the Sespe deposits to within at least ten feet of the stratigraphic horizon represented by Locality 150.

**Family Amynodontidae: Amynodontopsis bodei Stock**

Several fragmentary jaws and teeth from Locality 150 and a lower jaw, No. 2000 C.I.T. Coll., Plate 2, figures 1, 1*a* are referred to this species. The latter specimen comes from Locality 147 C.I.T., 160 feet higher stratigraphically than Locality 150 in the Sespe deposits, north of the Simi Valley, California. A lower canine tooth, No. 2001, from Locality 150, is shown in Plate 2, figure 2.

The lower jaw, No. 2000, represents an individual larger than *Amynodon erectus* (No. 11453 Y.P.M.) and *A. advenus* (No. 3102 C.M.) and resembles in this character *A. intermedius*. Of the two incisors present in the Sespe specimen, the lateral is decidedly larger than the medial one. Resemblance is shown to *Metamynodon* in this disparity in size. The occlusal surface of each incisor presents a rather curious pattern, suggestive however of the type of surface seen in comparable teeth of *Amynodon*. Extending back from the labial to the lingual border of each tooth is a ridge and on either side of this the enamel surface is indented or cupped. While only two incisors are preserved, an additional tooth is at least suggested by a faint furrowing of the alveolar wall of the symphyseal region in advance of the socket for the canine. If the tooth were present, its size, in relation to that of the incisor next to it, must have been small.

The canine possesses a short crown and long root. A well-marked furrow extends the length of one surface of the root. In addition to the worn posterior surface, a small surface of abrasion is present along the anterior edge at the base of the crown.

Only two lower premolars are present, and by far the larger portion of the wearing surface of the cheek-tooth series is furnished by the molars. The latter are individually longer in relation to their width than are the comparable teeth in *Amynodon*. The molars are less hypsodont than in *Metamynodon* and exhibit on their outer faces, as in *Amynodon*, the longitudinal groove which marks the union of protolophid and hypolophid. The furrows, however, are not so deep as in *Amynodon*.

The symphyseal region of the jaw is narrow and the dorsal surface

posteriorly is deeply sunk. A posterior mental foramen is situated below the anterior root of $P\overline{3}$.

*Amynodontopsis*, described[4] on the basis of a skull and dentition from the Sespe (Locality 150), is not so far advanced beyond *Amynodon* as is the Oligocene *Metamynodon*. Additional characters which point toward progressive development of the Sespe genus are presented by the lower jaw and dentition. These are the anteroposterior lengthening of the molar dentition with notable increase in extent of molar surface in relation to premolar surface, tendency toward greater flattening of external walls of cheek-teeth with fainter development of groove between protolophid and hypolophid, and in the characters of the incisors.

MEASUREMENTS (IN MILLIMETERS) OF NO. 2000 C.I.T.

| | | | |
|---|---|---|---|
| Length of symphysis | | | 108.5 |
| Depth of ramus at anterior end of $M\overline{1}$ | | | 65.5 |
| Width of ramus below $M\overline{2}$ | | | 24.6 |
| Length of diastema between $C$ and $P\overline{3}$ | | | *a*48.5 |
| Length from anterior end of $C$ alveolus to posterior end of $M\overline{3}$ | | | *a*230 |
| Length of premolar series ($P\overline{3}$–$P\overline{4}$) through middle | | | 39.8 |
| Length of molar series through middle | | | 119.5 |
| Medial incisor, transverse diameter | 9.9; | anteroposterior diameter | 10.6 |
| Lateral incisor, transverse diameter | 14.9; | anteroposterior diameter | 15.7 |
| $P\overline{3}$, transverse diameter | 12.8; | anteroposterior diameter | 17.7 |
| $P\overline{4}$, greatest transverse diameter | 17.1; | anteroposterior diameter | 23.1 |
| $M\overline{1}$, greatest transverse diameter | 20.7; | anteroposterior diameter through middle | 31.5 |
| $M\overline{2}$, greatest transverse diameter | 23 ; | anteroposterior diameter through middle | 40 |
| $M\overline{3}$, greatest transverse diameter | 21.4; | anteroposterior diameter through middle | 44.3 |

*a* Approximate.

[1] C. Stock, *Proc. Nat. Acad. Sci.*, **19**, 762–767 (1933); **21**, 456–462 (1935).
[2] H. E. Wood, 2nd, *Bull. Amer. Pale.*, **13**, No. 50, 16–19, Table II (1927).
[3] H. E. Wood, 2nd, *Ibid.*, pl. 3, Fig. 9 (1927).
[4] C. Stock, *Proc. Nat. Acad. Sci.*, **19**, 762–767 (1933).

# EXTRATERRESTRIAL EFFECTS OF COSMIC RAYS

By F. Zwicky

Norman Bridge Laboratory of Physics, California Institute of Technology

Communicated March 25, 1936

*A. Introduction.*—So far only interactions of cosmic rays with *terrestrial* matter have been investigated. If cosmic rays are present in interstellar and intergalactic space, they produce physical changes on extraterrestrial objects also. The question therefore arises if such changes are observable. The following considerations make it probable that this question is to be answered in the affirmative.

*B. Radiations in Interstellar Space.*—The following corpuscular and electromagnetic rays are traveling continually through interstellar space.

(1) *Light* from the *stars.*—Its total intensity $I_L$ in an average point in our galaxy is estimated to be of the order

$$I_L = 5 \times 10^{-3} \text{ ergs/cm.}^2 \text{ sec.} \quad (1)$$

(2) *Atomic rays* are ejected from novae, Wolf-Rayet stars, etc. These rays consist of atoms, ions and molecules, whose greatest individual kinetic energies are at least as high as $10^7$ e.-v. The total transport of energy by atomic rays probably corresponds to an intensity $I_A$ not smaller than

$$I_A = 10^{-4} \text{ ergs/cm.}^2 \text{ sec.} \quad (2)$$

Atomic rays, so far, have not been considered in astrophysics. More information about them will be given in another publication.

(3) *Cosmic Rays.*—The problem of their composition and their mode of creation has not yet been solved. Individual cosmic ray particles may possess energies of $10^9$ e.-v. and more. The total energy which is transported through space by the known high energy components of the cosmic rays corresponds to an intensity $I_C$ not less than

$$I_C = 3 \times 10^{-3} \text{ ergs/cm.}^2 \text{ sec.} \quad (3)$$

There can be no doubt that the gap between energies of $10^7$ e.-v. to $10^9$ e.-v. is filled by a variety of rays which may be either low energy cosmic rays or atomic rays more energetic than those which are emitted from bright common novae. Future observations on novae whose absolute visual brightness is greater than $M_{vis} = -9$ promise to clear up this point.

It is also desirable to analyse the light from the non-polar as well as from the polar aurorae in order to obtain more information about all those components of interstellar rays which do not penetrate deeply into the earth's atmosphere.

At the present time it is difficult to estimate the *total energy* transported

by light, atomic rays and cosmic rays. The composite intensity $I$ of all these rays is certainly greater than

$$I = 10^{-2} \text{ ergs/cm.}^2 \text{ sec.} \quad (4)$$

Judging from the high surface brightness (ca. $5 \times 10^{-3}$ ergs/cm.$^2$ sec.) of the non-polar aurora it would seem that the intensity $I$ is even considerably greater than estimated in (4). In the following we shall assume that in an average point in interstellar space of our own galaxy the total energy transport $I$ is of the order

$$I = 0.1 \text{ ergs/cm.}^2 \text{ sec.} \quad (5)$$

It must be kept in mind that for many purposes the effects of various rays which contribute to the intensity $I$ must be considered separately. For other purposes only the over-all value of $I$ is important.

*C. Excitation of Emission Lines by Cosmic Rays and Atomic Rays.*— Energetic corpuscles and photons which travel through interstellar space are capable of exciting emission lines which correspond to high stages of excitation. A closer scrutiny reveals that extended gaseous masses represent the most favorable objects for the production of emission lines by cosmic rays. Our attention therefore turns to the tenuous atmospheres of giant stars, interstellar gas clouds, comet tails, the upper parts of the earth's atmosphere and the like.

For the sake of illustration we shall confine the discussion to long period variable stars. These stars are known to be giants of very low surface temperature (ca 2000°K.). The fact that the spectra of long period variables show the Balmer lines in emission with high intensity has long constituted a puzzling problem. The energy necessary to excite the Balmer series is greater than 13.5 e.-v. which can only be furnished by *ultra-violet light* of wave-lengths shorter than $\lambda = 911$ Å or by *energetic particles.* Existing theories which try to explain the presence of either one of these two agents in the atmospheres of low temperature stars are not very convincing. As a new alternative I therefore propose to investigate the possibility that the emission lines in Mira stars are produced by energetic interstellar rays.

The interferometric diameters of long period variables are of the order of $2 \times 10^{13}$ cm. (Mira ceti). The actual atmospheres of these stars may, however, well extend to distances of $R = 10^{14}$ cm. from the center of the star, because the interferometer probably measures only the diameter of a dense and luminous central part of the star. The place of origin of the emission lines is most likely to be sought for in the outlying and extended regions of the star's atmosphere. If we assume that all of the interstellar rays which penetrate this atmosphere are absorbed, the total energy $E$ available per second is of the order

$$E = 4\pi R^2 I = 4\pi \times 10^{28} I \tag{6}$$

or

$$E \sim 10^{28} \text{ ergs/sec.} \tag{6a}$$

Since the star's atmosphere is presumably made up mostly of hydrogen, most of the absorbed energy will reappear in the form of hydrogen emission lines. The Lyman lines, being resonance lines of the normal state of hydrogen, cannot easily escape from the star's atmosphere. After emission they will be absorbed, reëmitted and reabsorbed again and again. This chain of events leads finally to a high concentration of monochromatic light whose wave-length is that of the first Lyman line $\lambda = 1215.7$ Å. The net result is that essentially only the Balmer lines escape from the star with high intensity. The arguments here given are essentially the same on which *Zanstra* based his successful interpretation of the spectra of planetary nebulae.

The high intensity of the resonance radiation $\lambda = 1215.7$ Å produces a high concentration of excited hydrogen atoms. These excited atoms may absorb radiation from the star itself as well as interstellar light. In this way the emission of the Balmer lines will be intensified.

Inasmuch as the total light emitted on the average from long period variables in the form of Balmer lines probably does not exceed $10^{31}$ ergs/sec., the above considerations suggest that interstellar rays may indeed be the prime cause for the presence of high excitation emission lines in the spectra of low temperature giant stars. It will be of interest to carry out the theory in more detail as soon as more information is available regarding the possible values of the intensities of the different radiations which are present in interstellar space.

In addition it must be investigated what part of the hydrogen resonance radiation is transformed into kinetic energy and is lost in the form of heat by absorption in various types of molecules and subsequent impacts of the second kind. These impacts become the more important the denser the atmosphere. If on the other hand the density of the star's atmosphere should be too low, the Lyman lines can themselves directly escape from the star. The efficiency of production of the hydrogen emission lines will therefore be highest for some definite density distribution and extension of a star's atmosphere.

As mentioned before, the presence of emission lines in the spectra of comet tails, of the night sky and of certain gas clouds in the Milky Way and in extragalactic nebulae represent additional phenomena which may possibly be caused by atomic rays and cosmic rays.

*D. Transfer of Mechanical Momentum from Atomic Rays and Cosmic Rays to Stellar and Interstellar Matter.*—For photons and particles whose

velocity is nearly equal to that of light, that is, for cosmic rays, the specific transport of momentum (per unit time in the direction normal to a given unit area) is

$$j_c = I_c/6c. \tag{7}$$

For particles whose velocity $v$ is appreciably smaller than $c$, that is, for most atomic rays, we have

$$j_A = I_A/3v. \tag{8}$$

If all of the radiation represented by the intensity $I$ is absorbed, the pressure exerted on the absorbing medium is $p = j$. With the values of $I_c$ and $I_A$ adopted in (2) and (3) we obtain $p_c = 1.66 \times 10^{-14}$ dynes/cm.$^2$ and $p_A = 3.3 \times 10^{-10}/v'$ dynes/cm.$^2$ where $v'$ is measured in kilometers per second.

Since atomic rays are more easily absorbed than cosmic rays they will be relatively most effective in imparting momentum to tenuous matter through which they pass. The possibility therefore suggests itself that the mechanical behavior of extended gas and dust clouds in interstellar space is not alone determined by gravitational forces and the pressure of light, but that the actions of interstellar corpuscular rays must be considered as well. The sharp outlines of such objects as the network nebula in Cygnus and the distinct contours of many interstellar dust clouds can perhaps be explained on this basis. In contradistinction to atomic rays, cosmic rays rather tend to disperse in course of time all local accumulations of interstellar gases.

*E. Absorption of Cosmic Rays in the Milky Way.*—Because of the eccentric position of the earth relative to the center of the Milky Way the absorption of cosmic rays on their passage through interstellar gas and dust clouds should produce a slight directional asymmetry of the intensity distribution of cosmic rays. This problem has recently been treated in these PROCEEDINGS.[1]

*F. Secondaries from Cosmic Rays and Atomic Rays.*—Individual particles of cosmic rays and atomic rays in course of time lose energy and are finally stopped entirely through formation of secondaries. The "mean free path" of cosmic rays is in general large compared with the linear dimensions of a galaxy ($10^{23}$ cm.) whereas the mean free path of atomic rays inside of our Milky Way is probably less than one light-year. Cosmic rays and atomic rays are therefore both ejected and re-collected by stellar system, thereby completing a cycle which may play an essential rôle in the evolution of stars and galaxies.

Finally we mention that interesting possibilities would arise regarding the origin of the hardest components of cosmic rays if there existed parts of the universe in which the carriers of positive and negative electricity were reversed compared with the elementary particles known to us. Mutual

annihilation of oppositely charged particles might then be expected to take place in certain regions of space.

*G. Final Remarks.*—The above considerations should be regarded as preliminary *programmatic* suggestions concerning theoretically possible extraterrestrial effects of cosmic rays and atomic rays. In order to proceed efficiently with this program it will be necessary to collect more data on the ejection of corpuscular rays from novae. As Baade and Zwicky[2] have suggested, the study of *super-novae* in particular promises to furnish the key to many phenomena related to interstellar rays, as well as to other fundamental problems of astrophysics.

[1] F. Zwicky, These PROCEEDINGS, **22,** 182–187 (1936).

[2] W. Baade and F. Zwicky, These PROCEEDINGS, **20,** 254 and 259 (1934); *Phys. Rev.*, **46,** 76 (1934).

---

## *THE EFFECT OF TEMPERATURE ON THE RESPIRATION OF THE EARTHWORM*

BY C. M. POMERAT AND M. X. ZARROW

THE BIOLOGICAL LABORATORIES, CLARK UNIVERSITY

Communicated March 21, 1936

A large number of studies dealing with the effect of temperature on the respiration of poikilothermous animals have consistently shown an increase in the rate of gas exchange with a rise in temperature. This relationship has been described by the application of the Arrhenius equation (cf. Crozier[1,2]) which states that the velocity of essentially irreversible reactions is proportional to the exponential of $-\mu/RT$. The equation is written

$$V = e^{-\mu/RT} + C$$

where $V$ is the velocity of the reaction being measured, $e$ is the base of the natural logarithms, $R$ is the gas constant, $T$ is the absolute temperature, $C$ is a constant of integration and $\mu$ is the critical thermal increment. In plotting the data and determining values for the constant $\mu$, the form of the equation used is

$$\log K_2 - \log K_1 = \frac{\mu}{2.3R}\left(\frac{1}{T_1} - \frac{1}{T_2}\right)$$

where $K_1$ and $K_2$ are the rates of respiration at the absolute temperatures $T_1$ and $T_2$. Temperature characteristics have repeatedly been found to yield values of approximately 11,000 and 16,000 calories for respiratory processes.

Vernon[3] reported a possible exception to this rule in the earthworm whose respiratory rate remained practically constant between the temperatures of 10° and 22.5°C. These experiments have been interpreted as showing "that even these lowly forms must have some mechanism for restraining oxidation" (Mitchell[4]). A nervous mechanism for this regulation was suggested by Vernon. Krogh[5] has pointed out that Konopacki[6] failed to find any indication of constant metabolism in the earthworm at temperatures between 10° and 20°C. and concluded that Vernon's result might have been accidental and due to imperfect control over conditions other than temperature.

Respiratory measurements have been made in this laboratory on the earthworm, *Lumbricus terrestris*. Normal, decerebrate, and animals from which the subesophageal ganglion had been removed were studied. All animals were collected at the same time (September 27, 1935) and were kept under identical conditions in jars containing wet filter paper. Data were obtained during a period of two weeks beginning with the second day of laboratory life. No shift in the relative values for respiration at different temperatures was observed during this period although absolute values tended to be lowered with time. This was well correlated with the decrease in the volume of the animals.

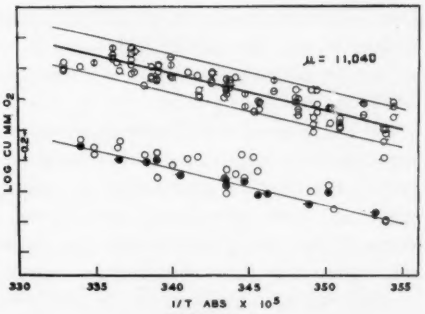



FIGURE 1

The rate of $O_2$ consumption as a function of temperature in the earthworm. The upper group of values was obtained from normal animals. In the lower group open circles refer to decerebrates and solid circles to animals whose subesophageal ganglia had been removed.

Measurements were made by the direct method of Warburg using a constant-volume type of respirometer and a special vessel of approximately 30-cc. capacity. Constant temperatures were maintained by means of a thermostat.

The rate of $O_2$ consumption was determined for temperatures ranging from 9° to 27°C. Measurements were made for both ascending and descending series of temperatures. Figure 1 shows the relation between the logarithm of the rate of $O_2$ consumption in cu. mm. and the reciprocal of the absolute experimental temperature. Single individuals are indicated by various symbols. The upper band of values represents a composite plot of data made in such a way as to superpose individual curves. The same abscissas were used but the differences in the absolute values of the ordi-

nates were eliminated. The slope of the line drawn through these points yielded a temperature characteristic of 11,040 ± calories. The data plotted beneath the normal composite represent decerebrates (open circles) and animals from which the subesophageal ganglion had been removed (solid black circles). The ordinates have been multiplied by a suitable constant to separate the two bands of data. A line exactly parallel to that determined for normal individuals was drawn through the points obtained from operated animals. No significant shift in the value of $\mu$ seems indicated. Greater variability in $O_2$ consumption occurred in the decerebrates while animals without subesophageal ganglia showed remarkably regular respiration. These facts are correlated with increased activity in the former and decreased activity in the latter group. The absolute values obtained in all determinations reported in this study varied between 25 and 250 cu. mm. of $O_2$ per half-hour period.

On the basis of these experiments it is concluded that the relation existing between temperature and respiratory rate in the earthworm is similar to that reported for other poikilothermous animals which have been studied under adequate experimental conditions. It is, therefore, not necessary to postulate a special mechanism for restraining oxidations in this species. The removal of the cerebral and subesophageal ganglia has not been found to cause shifts in the value of the $\mu$ obtained in the normal animal.

[1] Crozier, W. J., *Jour. Gen. Physiol.*, **9**, 531 (1925–26).
[2] Crozier, W. J., *Ibid.*, **7**, 189 (1924–25).
[3] Vernon, H. M., *Jour. Physiol.*, **19**, 18 (1897).
[4] Mitchell, P. H., *A Textbook of General Physiology*, McGraw-Hill Co., 1932.
[5] Krogh, A., *The Respiratory Exchange of Animals and Man*, Longmans, Green & Co., 1916.
[6] Konopacki, M., *Bull. de l'Acad. de Sci. de Cracvie, M. n. Kl.*, 357 (1907).

---

## *THE CHANGE IN THE RESPONSE OF AVENA-COLEOPTILES TO GROWTH REGULATORS PRODUCED BY AGING*

By H. G. du Buy

The Biological Laboratories, Harvard University

Communicated April 9, 1936

The curvature of Avena-coleoptiles and other plants is often taken as a quantitative index of the action of growth regulators and of other processes. The variability of the responses has made it necessary to define exactly the conditions under which the plants must be cultivated in order to make this method useful. It is, however, not well known why these conditions

must be fulfilled, and what changes cause variation in the curvatures. Failure to realize these sources of variation has occasionally led to error.

The present paper shows the amounts of change produced by aging, with consideration of the causes of these changes. Finally, the conditions which must be met in order that the curvature may be used analytically are discussed.

Decapitated Avena-coleoptiles of different stages of development (ages) show different amounts of curvature when the same amount of auxin A is applied unilaterally: i.e., their response to auxin expressed in elongation (Reaktionsvermögen) is decreased. Plants which have been illuminated show the same decrease in response although the phenomena are more

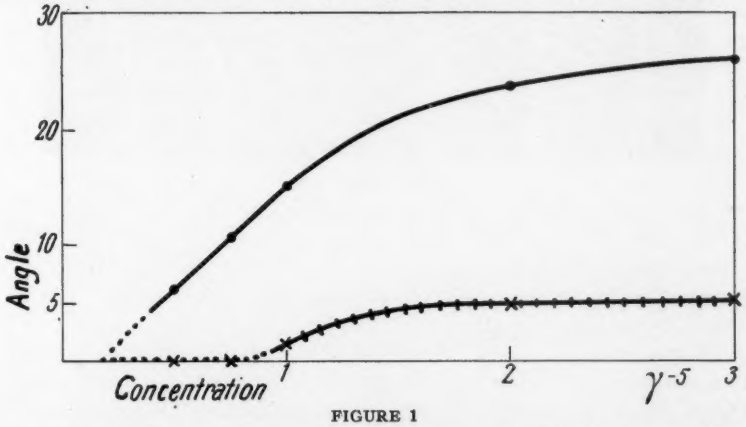



FIGURE 1

Upper curve: Reaction of young coleoptiles. Lower curve: Reaction of old coleoptiles. Each point of the curve represents the average curvature of a series of 12 plants.

complex here. Figure 1 shows the angle of curvature of decapitated plants as a function of the concentration of auxin, the upper curve being for young plants (3 days) and the lower one for older plants (6 days). It is seen that older plants become completely insensitive to the auxin when it is applied below a certain concentration; in other words, the threshold of response is seen to increase with age. Figure 2 shows the difference in response of young and old plants on $2.10^{-5}$ γ auxin.

This age effect is due to a change in elasticity and irreversible extensibility of the cell wall as shown by du Buy (1933, p. 849 ff).

Other experiments (l.c. p. 851 ff) have shown that the production of growth regulators decreases with increasing age. Therefore, at a certain age *the concentration of auxin produced in the plant falls under the threshold value* mentioned above *so that the plant stops growing altogether*, even shortens under certain conditions.

However, if these plants are placed horizontally or if they are illuminated unilaterally for a sufficient length of time, the existing growth regulator migrates to the lower or the non-illuminated side, so that *there* the threshold will be exceeded. *A curvature occurs although normal growth no longer can be*

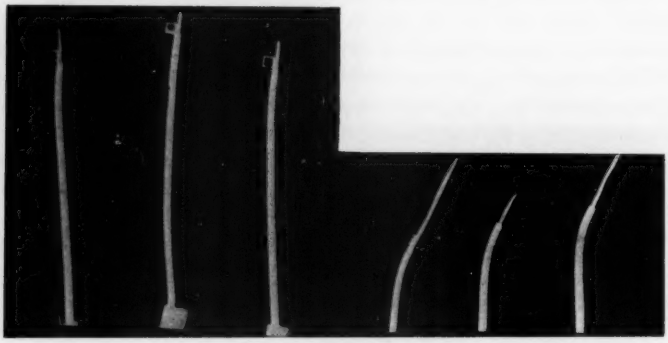


FIGURE 2

Left: Reaction of old coleoptiles. Right: Reaction of young coleoptiles on the same amount of auxin.

*measured.* This explains the results obtained by some investigators (Beyer, 1932).

Du Buy and Nuernbergk (1929) have shown that there is always lateral transport of auxin if there is a difference in concentration of growth regulator between both sides of the plant and is demonstrated clearly in figure 3

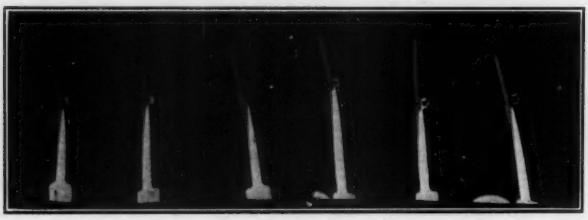


FIGURE 3

Shadowgraph of the same coleoptiles before and after their reaction on growth regulators.

by the change in place of the agarblock which contains the auxin. In this case a shadowgraph is made on the same paper at the beginning and at the end of the experiment. Both sides have grown considerably although the curvatures are small. In other experiments the curvatures can be greater

after application of the same amount of auxin: this is the so-called standard-phenomenon (Kögl and Haagen Smit, 1933). The lateral transport is always smaller in this case. The amount of lateral transport depends on the concentration of growth regulator and also on the conditions of the transporting and reacting cells which are changed by age, light, nourishment, etc.

We can conclude from these experiments that the auxin concentration which just causes a visible curvature ("threshold-concentration") varies with the conditions of the reacting cells. Other experiments have shown that it varies also with the properties of the organ (thickness, speed of transport, etc.). It is, however, common in the literature (e.g., Van Overbeek, 1933) to assume that for different conditions of plants the same threshold can be obtained which means that the two curves of figure 1 would intersect the horizontal axis at the same point which is obviously not the case. An extrapolation of the lower curve to this common intersection in order to obtain an *S*-curve (dotted line) as suggested by other investigators is therefore incorrect.

*Summary.*—With increasing age the response of Avena-coleoptiles to auxin decreases for all concentrations of this growth regulator. The threshold concentration for curvature increases with the age of the plants and not in the shape of an *S*-curve.

It is only in a certain stage of their development that the plants are sensitive to auxin. If these substances are applied before or after this period of sensitivity the response decreases or disappears totally. The duration of this period can be changed. This shows clearly that the substances act as a regulator. The value of curvature as a quantitative index under the above-mentioned conditions is discussed.

Beyer, A., "Untersuchungen zur Theorie der pflanzlichen Tropismen. Die Beziehungen zwischen geotropischer Krümmung und Wachstum bei Dikotylen," *Planta*, **18**, 509 (1932).

Buy, H. G. du, "Ueber Wachstum und Phototropismus von Avena sativa," *Rec. trav. bot. néerl.*, **30**. Diss. Utrecht (1933). (Also, *Ber. der deutschen bot. Ges.*, **52** (1934).)

Buy, H. G. du, and Nuernbergk, E., "Weitere Untersuchungen über den Einfluss des Lichtes auf das Wachstum von Koleoptile und Mesokotyl bei Avena sativa. II. Mitteilung," *Proc. Kon. Akad. Wetensch. Amsterdam*, **32**, 808 (1929).

Buy, H. G. du, and Nuernbergk, E., "Phototropismus und Wachstum der Pflanzen," *Ergebn. der Biologie*, **9, 10, 12.** Berlin. (1932, 1934, 1935).

Kögl, F., and Haagen Smit, A., "Ueber Auxine," *Z. angew. Chemie*, **46**, 469 (1933).

Overbeek, J. van, "Wuchsstoff, Lichtwachstumsreaktionen und Phototropismus bei Raphanus," *Rec. trav. bot. néerl*, **30**, 537 (1933).

# THE SEXUAL AND REPRODUCTIVE CYCLES OF CHIMPANZEE[1]

BY ROBERT M. YERKES AND JAMES H. ELDER

YALE LABORATORIES OF PRIMATE BIOLOGY

Read before the Academy, April 28, 1936

Present knowledge of sexual and reproductive processes in the chimpanzee is unsatisfactory because incomplete and unreliable. It rests principally upon observation—too often incidental, brief and inexpert—of an occasional captive specimen of inadequately known life history, characteristics and status, in relatively unnatural and reproductively unfavorable conditions. Heretofore, because of lack of standards of judgment based upon statistically valid data, it has been impossible to classify and evaluate case studies with respect to normality and typicalness. Manifestly descriptions based upon observation of a single case, however fully studied, may be grossly misleading. There is urgent need for a body of comparable data which shall supply trustworthy descriptions for genus and species, supplementary to those now available for individuals.

As fruit of the opportunity created by Yale Laboratories of Primate Biology, it is now possible to present data in partial satisfaction of this need. The materials upon which this summary report is based include the results of systematic study of fourteen sexually mature females over periods varying from months to several years. In a few instances the history of an individual has been followed from childhood to reproductive maturity. Our protocols include records of some two hundred complete sexual cycles observed day by day; descriptions of about a score of complete reproductive cycles, and of about half as many incomplete or interrupted cycles; and, finally, of data of behavior and reproductive outcome for somewhat more than five hundred controlled matings. The studies have been conducted in a situation created for reliability of observation and satisfactoriness of experimentation.

Lest we help to perpetuate some common misunderstandings, we wish to offer certain definitions. The expression "sexual cycle" we shall use consistently to designate that periodical succession of events in the reproductive system of the female which may include some or all of the following phenomena: menstrual bleeding, genital swelling, oestrus and ovulation. We state the matter thus because the sexual cycle, in our usage, may or may not exhibit all of these events and should not be confused or identified with any one of them. Genital swelling, for example, may occur apart from menstrual bleeding, and it appears that in an otherwise complete sexual cycle ovulation may be absent. The term "sexual cycle" stands then as inclusive and enables us to designate as phases such phenomena as menstruation, swelling, oestrus and ovulation.

For that sequence of biological events which results in the reproduction of the parents, we use the term "reproductive cycle." This cycle is necessarily inclusive of the sexual cycle. It comprehends ovulation, mating, fertilization, gestation, parturition, lactation and care of progeny. Our usage is especially remarked because the succession of happenings which typically results in reproduction is often designated the "sexual cycle," or the "generative cycle."

*The Sexual Cycle.*—As the sexual cycle of chimpanzee may readily be visualized in its essential phases, we shall use two diagrams to facilitate description. The one represents the succession and temporal relations of events in what we consider a typical multiparous female whose sexual and reproductive history is known for several years; and the other similarly represents the facts for the genus Pan, in that data for ten presumptively normal, sexually mature females have been combined.

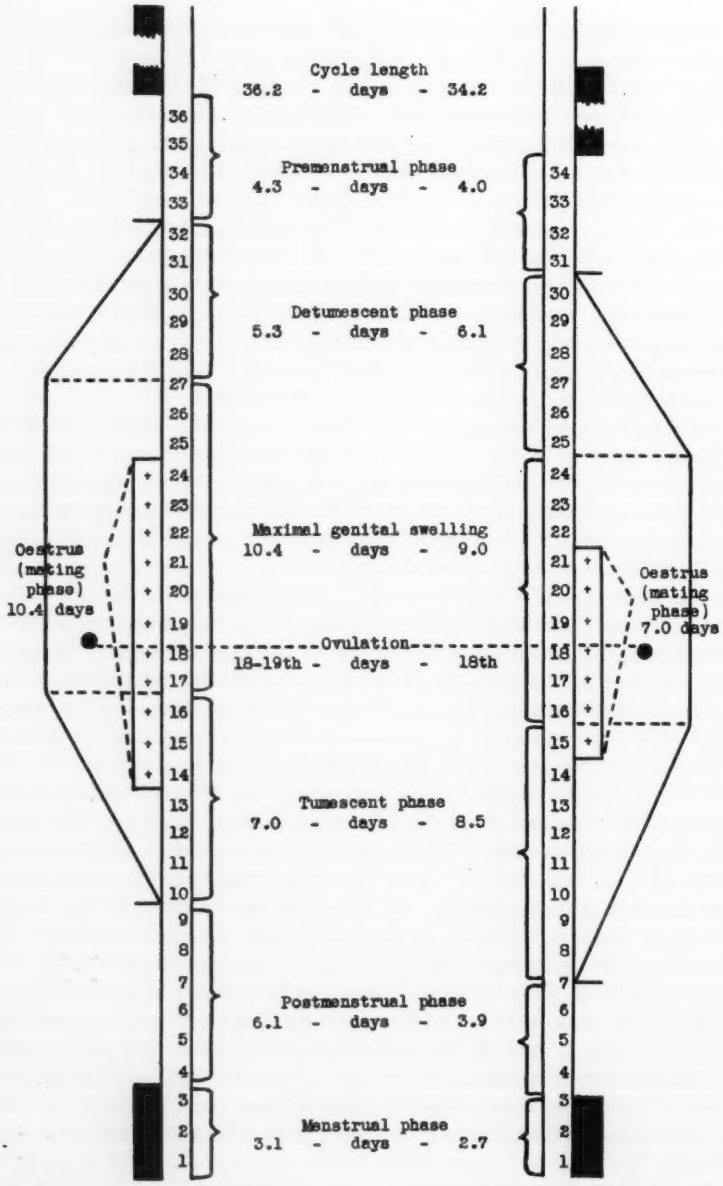
Of the fourteen females previously referred to as subjects of observation, four were omitted in our quantitative treatment of results for the following reasons: one case, because of paucity of observations and previous description by Tinklepaugh;[2] one case, because recently matured sexually and never impregnated; and two cases, because obviously abnormal as well as atypical sexually and reproductively.

Although, logically, menstrual bleeding may belong at the end of the cycle, we have placed it at the beginning in our diagrams because it is found more convenient to date the train of events from the appearance of bleeding than from any less precisely and readily determinable event.

In the accompanying figure the right diagram represents the sexual cycle of the individual Mona as plotted from the averages for the data of fifteen complete cycles which were observed between January 15, 1932, and October 14, 1935. During the remainder of the interval between these dates Mona was either pregnant or lactating. Determination of the period of sexual receptivity (oestrus) in this individual is based upon sixty-three controlled matings with the mature breeding male Pan. It appears that female receptivity, as tested by controlled mating, continues in this female for about seven days during the maximal phase of genital swelling. We are aware that copulation is not a wholly satisfactory test of oestrus or female receptivity, since it depends upon two individuals either of which on a given date may be negative sexually toward the other. The indicated date of ovulation, the 18th day of the cycle, is based upon two observations in which a controlled mating on that day of the cycle resulted in pregnancy. Naturally we should not venture to present these observations as statistically significant were they not in agreement with results obtained with other females.

The reader will note that the average duration of the sexual cycle for Mona is 34.2 days; that menstrual bleeding lasts about 3 days, and that

NORMS FOR THE SEXUAL CYCLE OF CHIMPANZEE




Norms for the genus Pan, based on data for 10 individuals (158 cycles)

Norms for the individual Mona, based on data for 15 cycles

the succeeding phases of the cycle are describable thus: duration of the postmenstrual phase, 4 days; of the swelling phase (inclusive of tumescence, maximal genital swelling and detumescence), 24 days; of the premenstrual phase, 4 days.

The diagram for the genus Pan, presented at the left of the accompanying figure, is constructed according to the same principles as that for the individual Mona. It is based upon the averages for 10 individuals for whom a total of 158 records of complete sexual cycles were available. Since this appears to be a reasonably adequate number of observations as predictive basis, we present the following averages as tentative norms for chimpanzee. For the duration of the sexual cycle the norm is 36.2 days, or approximately five weeks, as compared with the four-week cycle of man. The range of average cycle durations for the 10 individuals included in our assemblage is 33.8 to 44.4 days. The shortest individual cycle observed is 29 days. The several phases of the sexual cycle exhibit the following durations: the menstrual phase, 3.1 days; the postmenstrual phase (sexual skin detumescent), 6.1 days; the period of gradual enlargement of sexual skin, 7 days; the continuance of maximal genital swelling, 10.4 days; the length of oestrus or sexual receptivity, 10.4 days; the appearance of ovulation, 18–19th day of cycle; the period of detumescence, 5.3 days; the premenstrual phase, 4.3 days. It appears from our data that ovulation may occur at the mid-point of the sexual cycle or slightly thereafter, and that for a few days both before and after that event the female may be in oestrual phase and willing to mate. About the conditions, relations, frequency and dates of ovulation we know least at present, for as yet we lack methods of direct observation and are dependent wholly upon the reproductive results of controlled matings.

In this brief report we may not attempt description of variations in the sexual cycle; they are numerous and extreme from individual to individual and at different periods in the sexual life of a given female. Consequently only the experienced observer, or he who has such data as we are now attempting to present to guide him, may hope to avoid inadequacy of observation, confusion and misinterpretation.

We do not wish the reader to infer that we consider our observations relative to the relations and durations of phases of the sexual cycle in chimpanzee adequate: they are only relatively and tentatively so. But whatever the reliability of our judgments concerning the normality and typicalness of our subjects, we have the satisfaction of presenting the first norms for the sexual cycle of chimpanzee.

*The Reproductive Cycle.*—In supplementation of our summary account of the sexual cycle, we shall now offer partial description of the reproductive cycle. To begin with ovulational process, we have not as yet attempted experimental control, but by the procedure of controlled mating

we have observed nine instances of pregnancy which followed ovulation as early as the 16th day of the sexual cycle and as late as the 24th day. There is no instance in which it anticipated the middle of the cycle. For observational verification we propose the hypothesis that whereas the interval between the beginning of menstruation and the occurrence of ovulation varies with cycle length, that between ovulation and the end of the sexual cycle is constant. Our data suggest also the hypothesis that the distribution of individuals in accordance with cycle length is bimodal: presenting a short cycle group with mode at 35–36 days and a long cycle group with mode above 40 days. For the latter group ovulation would occur relatively later in the cycle than for the former.

We are not able to characterize the process of fertilization, nor to throw any important light on the frequency of sterility. Of the adult chimpanzees used—four males and fourteen females—all of the former and twelve of the latter have demonstrated fertility. Furthermore, it is indicated by experience in this colony that the chimpanzee may be bred and reared without peculiar difficulties in a laboratory colony and under such nutritional and general hygienic conditions as we have been able to maintain.

Mating behavior and female receptivity we have studied systematically under controlled conditions and therefore are able to present certain descriptive statements and to designate some at least of the important factors which condition mating, thus supplementing or correcting current knowledge and preparing the way for further observational and descriptive effort.[3] The surmise is current, even among those who are especially familiar with primate physiology, that chimpanzees may mate at any time in the sexual cycle and that relations between consorts typically continue during pregnancy. These facts have been asserted for various primate types, but our observations indicate that they do not apply to chimpanzee. As affirmation of the occurrence of oestrus and of the periodicity of female receptivity and mating in chimpanzee, we present the following summary statement.[4] It is based upon observational data obtained from between five and six hundred controlled matings which involved the use of all of our sexually mature subjects. We desire especially to stress the fact that each observation represents a definitely prearranged experiment, with two subjects of known history, sexual status and social relationship. Such conditions for the study of primate mating are unparalleled in the scientific study of the subject.

Normally and typically a multiparous female will mate with a male with whom she is intimately acquainted and on friendly or affectionate terms only during a few days of the period of genital swelling about and including the usual time of ovulation. At other times the female as a rule is ignored sexually by the male, or, if approached or solicited, she avoids him. Every

word in this statement is important, for social relationships as well as the physiological status of the consorts are important determinants of response. If either mate be recently matured sexually, relatively inexperienced in mating, or if the two be strangers, slightly acquainted, or hostile mutually or one to the other, mating may occur under any condition of sexual status and seemingly irrespective of biological appropriateness and reproductive requirement. Or, on the other hand, it may not occur even under entirely appropriate physiological conditions. Inasmuch as most reports concerning the mating behavior of chimpanzees have ignored the social factors to which we herein invite attention, it is inevitable that atypical behavior should have been mistaken for the biologically suitable, normal and typical. The pertinent literature and our extensive experience clearly enough indicate that unless the experimental conditions for the study of these phenomena simulate those of nature, unusual or atypical sexual behavior is to be expected. We should not maintain that our results are comparable with those which might be obtained in nature, but we have excellent reason to believe that they are far more trustworthy and also safely interpretable than any previously available.

Mating behavior varies extremely in and among individuals; that of a given female varies with different males. To one she may offer herself for sexual contact only for a few days about the middle of her sexual cycle; to another, she may present sexually over a more extended period or even at any time in the cycle. Such behavior often appears to be protective or precautionary and to serve the purpose of calming the male. In illustration of these statements we offer the following examples. Female Lia mated with male Pan on the 10th to the 19th cycle days; with male Bokar, practically throughout the cycle. Female Mona mated with male Pan from the 15th to the 20th day ordinarily, and with male Jack from the 8th to the 24th. In general, with intimate friendly consorts, the female takes the initiative in sexual approach, or at least definitely indicates receptiveness.

Relative to gestational process, our observations are unusual by reason of abundance and relative favorableness of circumstances. Reproductive accidents, due it seems chiefly to unfavorable physiological condition of the female, induced by dietary deficiencies or unsuitable hygiene or social factors, have been noted. We have records of ten miscarriages, early and late, over a period of six years, and in the same interval there have occurred sixteen normal births, of which thirteen were accepted as full-term.

The gestational period has been variously stated or estimated to be as short as five or as long as ten months; whether calendar or lunar ordinarily is not mentioned. In fifteen cases of our record[5] the interval between conception and parturition ranged from 202 to 261 days. These figures may be in error by one to six days, since ovulation and fertilization were not directly observed, but instead inferred from various observations of sexual

status and the results of controlled matings. The 202-day birth was clearly premature by at least 30 days. The shortest period for presumptively full-term birth in our array of cases is 216 days. The average and median gestational interval for full-term births is 236 days. This in our opinion, if erroneous, is too short. In the light of our observations we venture the statement that gestational interval of 236 days in chimpanzee is comparable with 266 days in man. There is a difference of approximately one month. If then we may speak of the human period as nine months, that of the chimpanzee should be stated as eight months.

It has not thus far appeared desirable to try to arrange in these laboratories for continuous observation of reproductive process including parturition. This is solely a matter of practical judgment. We have deemed it wiser to await favorable opportunity for economical use of our time than to institute such continuity of observation as would in every case guarantee completeness of description. It therefore happens that for fifteen cases of reproduction in this colony, parturition was observed in only three. From this small number of cases it would appear that orientation and presentation of the chimpanzee fetus more closely resemble those characteristic of man than of the monkey. We do not venture to particularize because our data are too few to justify it. Obviously the subject is one of special interest for further observation and for experimental inquiry with genetical factors and relationships especially in mind. The point which we desire to make is that resemblance in this particular phase of reproduction appears to be much closer between chimpanzee and man than between monkey and man or even monkey and chimpanzee.

Ordinarily parturitional process is prompt and easy in chimpanzee. Various reports of labor indicate continuance of contractions for from one-half hour to five hours prior to delivery. Normally the afterbirth is expelled shortly after the fetus, and we have observed no case of delay. Invariably the parturient female, if in normal condition, licks up from the floor of the cage available fluids—amniotic and blood—and partially or completely devours the afterbirth within a few hours. The latter, however, is not unexceptional, for in addition to those observed by us there are several cases of record in which it was left entirely untouched. Our observations strongly suggest that acceptance and appropriate care of the infant by the mother and normal lactation and nursing may occur whether or not the afterbirth is eaten. As we have no case in which the fluids were not taken by the female in appreciable quantity, we may not make the same statement regarding them.

Typically the experienced (multiparous) female accepts her infant as soon as born, takes it up, handles it more or less skillfully, gently and carefully, cleans it, permits or encourages it to cling to her coat and skin, and carries it about with her almost continuously. By contrast, the in-

experienced individual, when her first infant appears, may act as if bewildered, frightened, incredulous or mystified, and completely at a loss what to do. After some hesitation and delay, with sensory examination of the stranger, she may either take it up and care for it with a measure of appropriateness, or instead treat it as an object of curiosity, to be examined and tested in many ways, pushed or turned about, poked, sniffed at, even gently bitten, but not permitted to grasp and cling to her, much less to be hugged by her and carried about. In this latter case the infant has only slight chance of survival unless removed from the mother and bottle-fed.

[1] Acknowledgment is made to the Committee for Research in Problems of Sex, NATIONAL RESEARCH COUNCIL, for assistance at various times during the progress of this inquiry, and to Doctor O. L. Tinklepaugh, Doctor S. Zuckerman and Mr. M. I. Tomilin, who during periods of connection with these laboratories contributed to the development of techniques, modes of description and the accumulation of observations.

[2] O. L. Tinklepaugh, "Sex Cycles and Other Cyclic Phenomena in a Chimpanzee during Adolescence, Maturity and Pregnancy," *Jour. Morphol.*, **54**, 521–547 (1933).

[3] Complete accounts of this study of oestrus, receptivity and mating behavior in chimpanzee, and also of our observations on the sexual cycle, are in preparation for early publication.

[4] This is confirmatory of the findings of Tinklepaugh, op. cit., 529–530.

[5] J. H. Elder and R. M. Yerkes, "Chimpanzee Births in Captivity: a Typical Case History and Report of Sixteen Births," *Proc. Roy. Soc. Lond.* [Communicated.]

---

## *THE NEARCTIC DIRHININI AND EPITRANINI*

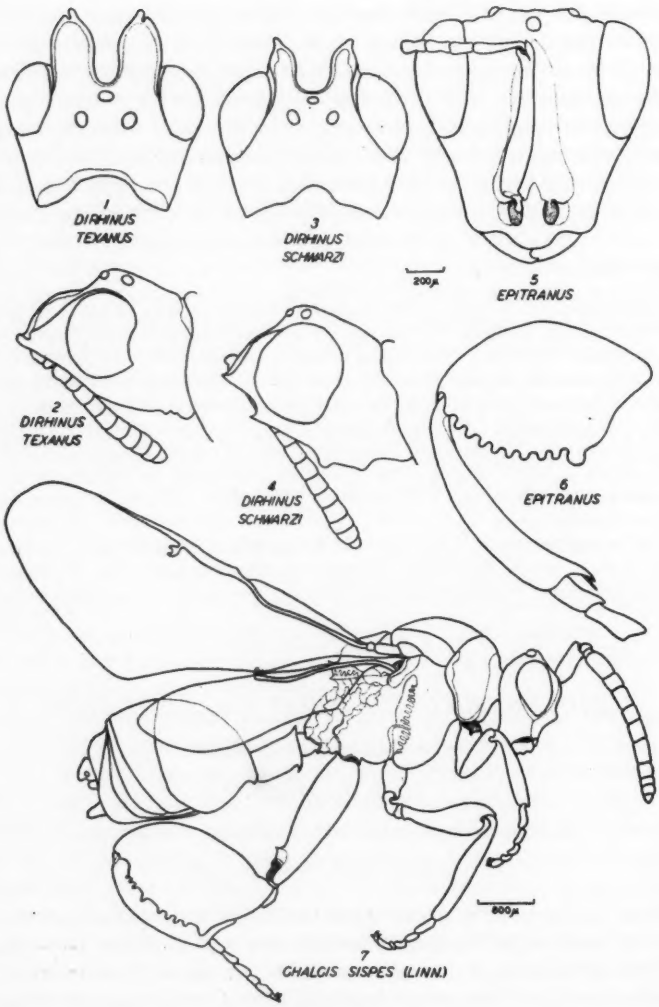
### (*Hymenoptera, Chalcididae*)

BY B. D. BURKS[1,2,3]

ILLINOIS STATE NATURAL HISTORY SURVEY, URBANA, ILL.

Communicated March 20, 1936

When a large number of specimens of Chalcid flies, collected over a wide area, or reared in rather large quantities, are studied, it becomes apparent that the genera, as at present established, are in an unsatisfactory and obscure condition. Almost all descriptions have been drawn up from small numbers of specimens, more often than not from single specimens, so that the published generic characterizations seldom provide for variations in structure. One of the most typical aspects of this group of parasitic insects is that almost all possible variations are to be encountered within a single species. The fact that new material very seldom exactly coincides with established types and published descriptions has led to the creation of an extremely large number of new genera and species. Some of these are un-




FIGURES 1–7

doubtedly valid, but many can be shown to be included in the range of normal variation of previously established genera and species.

As the members of the subfamily Chalcidinae are among the most bizarre and conspicuous Chalcids, they have been placed in a large number of small genera. The majority of these genera are either based on criteria

of only specific value, or the characters used intergrade with those of related genera to such an extent that it is practically impossible to make generic placements with certainty. The solution to this difficulty is to eliminate most of the genera, and include the species in a few large genera, which can then be defined simply and clearly.

Of the five tribes included in the Chalcidinae (the Chalcidini, Brachymerini, Haltichellini, Dirhinini and Epitranini) the last two will be considered here with reference to the nearctic fauna. The others will be treated elsewhere.

*Tribe Dirhinini.*—The typical form of a member of the Chalcidinae is shown in figure 7; the hind coxae are conspicuously enlarged and lengthened, and are round in section; the hind femora are enlarged and provided with a row of teeth on the lower side. The members of the tribe Dirhinini are easily to be recognized: the vertex of the head (Figs. 1, 2) is produced forward, laterad of the antennal grooves, into two prominent horns. This tribe in the nearctic region includes but a single genus.

Genus *Dirhinus* Dalman[4]

*Synonyms:* *Dirrhinus* Dalman[5]
*Eniaca* Kirby[6]
*Hontalia* Cameron[7]
*Pareniaca* Crawford[8]

*Dirhinus* Dalman was described for the species *D. excavatus* from Sierra Leone, and characterized as having 12 segmented antennae, which are inserted moderately far from the margin of the clypeus, and having the head deeply excavated mesally. As has been shown by Masi,[9] however, the number of antennal segments is actually 13, the usual number in the Chalcidinae. Kirby's *Eniaca* was described as having 13 segmented antennae, and differs in no other significant character from *Dirhinus*. Most subsequent workers, finding their material to have antennae with 13 segments, have referred their species to *Eniaca*. *Hontalia* Cameron, originally placed in the Aximini of the Pteromalinae, differs from *Dirhinus* and *Eniaca* only in that the frons is secondarily produced to form two smaller projections below the principal ones of the vertex (as in Fig. 4), and the female possesses a long, conspicuously-exserted ovipositor. The first character is certainly not of generic importance, and the second is not feasible as it occurs in only one sex. *Pareniaca* Crawford was set up on the same character as the first one used for *Hontalia:* the possession of secondary protuberances on the frons.

*Key to Nearctic Species.*—Anterior margin of vertex with a slight lateral notch (Fig. 1); frons not secondarily produced (Fig. 2)..... *texanus.*

Anterior margin of vertex not notched (Fig. 3); frons secondarily produced (Fig. 4)..... *schwarzi.*

*Dirhinus texanus* (Ashmead)[10]
(Figs. 1, 2)

*Host:* Dipterous larvae.
*Type:* Cat. No. 3325 U. S. N. M., female.
*Distribution:* Texas, Arizona, Kansas, Tennessee, Alabama, Mexico, Philippine Islands, China.

*Dirhinus schwarzi* (Crawford)[11]
(Figs. 3, 4)

*Host:* Dipterous larvae.
*Type:* Cat. No. 15547 U. S. N. M., female.
*Distribution:* Texas, Arizona, California.

*Tribe Epitranini.*—The tribe Epitranini has formerly been known as the Chalcitellini. The abdomen is petiolate (as in Fig. 7), the antennae are inserted just above the clypeal suture (Fig. 5), and the hind tibiae have a terminal spine-like projection (Fig. 6). This tribe in the nearctic region includes but a single genus.

Genus *Epitranus* Walker[12]

*Synonyms:* *Chalcitella* Westwood[13]
*Anacryptus* Kirby[14]
*Arretocera* Kirby[14]
*Platychalcis* Cameron[15]
*Arretoceroidella* Girault[16]
*Chalcitelloides* Girault[17]
*Arretoceroides* Girault[18]

The genus *Epitranus* Walker was misplaced by Ashmead[19] in his tribe Smicrini; the insertion of the antennae excludes it from that tribe and would refer it to his Chalcitellini. *Epitranus*, however, has priority over *Chalcitella* Westwood, and as the two are not generically distinct, *Chalcitella* must be sunk in synonymy, and the tribe be renamed. Kirby's *Anacryptus* and *Arretocera* were established for minor differences in the antennae, and cannot be held distinct from *Epitranus*. *Platychalcis* differs only in specific characters from *Epitranus*. The three genera proposed by Girault can safely be discarded as not sufficiently distinct, the only differences given by the author being small ones of the numbers of antennal segments.

*Key to Nearctic Species.*—Petiole smooth above, one-half the length of the abdomen..... *fulvescens.*

Petiole longitudinally striated above, two-thirds or more the length of the abdomen..... *castaneus.*

*Epitranus fulvescens* Walker[20]

*Host:* Unknown.

*Type:* In British Museum, female.

*Distribution:* West Indies, Brazil, and should occur in Florida.

*Epitranus castaneus* Cresson[21]

*Host:* Unknown.

*Type:* Cat. No. 1823.1 Phil. Acad. Nat. Sci., male.

*Distribution:* Cuba, and should occur in Florida.

[1] The study of the material in the U. S. National Museum was made possible by a grant from the Bache fund of the National Academy of Sciences.

[2] I am indebted to Mr. A. B. Gahan for considerable advice concerning genera the types of which are in Europe.

[3] Contribution No. 183 from the Entomological laboratory of the University of Illinois.

[4] Dalman, J. W., *Svensk. Vet.-Akad., Handl.*, **39**, 75 (1818).

[5] Dalman, J. W., *Analect. Entom.*, 29 (1823).

[6] Kirby, W. F., *Jour. Linn. Soc. Lond., Zoöl.*, **17**, 57 (1883).

[7] Cameron, P., *Biol. Cent.-Am., Hymen.*, **1**, 112 (1884).

[8] Crawford, J. C., *Proc. U. S. N. M.*, **45**, 312 (1913).

[9] Masi, L., *Ann. Mus. Civ. Stor. Nat. Genova*, **48**, 23 (1919).

[10] Ashmead, W. H., *Trans. Am. Ent. Soc.*, **23**, 217 (1896); Dalla-Torre, K. W. von, *Cat. Hymen.*, **5**, 368 (1898); Schmiedeknecht, O., *Gen. Ins.*, **97**, 67 (1909); Roberts, R. A., *Jour. Agr. Res.*, **50**, 490 (1935).

[11] Crawford, J. C., *Proc. U. S. N. M.*, **45**, 313 (1913).

[12] Walker, F., *Entom. Mag.*, **2**, 21 (1834).

[13] Westwood, J. O., *Proc. Zoöl. Soc. Lond.*, **3**, 70 (1835).

[14] Kirby, W. F., *Jour. Linn. Soc. Lond., Zoöl.*, **17**, 56 (1883).

[15] Cameron, P., *Invert. Pacif.*, **1**, 57 (1904).

[16] Girault, A. A., *Arch. Naturg.*, **78**, Abt. A, Heft 6, 66 (1913).

[17] Girault, A. A., *Ent. News*, **25**, 30 (1914).

[18] Girault, A. A., *Mem. Queensl. Mus.*, **4**, 351 (1915).

[19] Ashmead, W. H., *Mem. Carn. Mus.*, **1**, 251 (1904).

[20] Walker, F., *Ent. Mag.*, **2**, 26 (1834); Kirby, W. F., *Jour. Linn. Soc. Lond., Zoöl.*, **17**, 55 (1883); Dalla-Torre, K. W. von, *Cat. Hymen.*, **5**, 383 (1898); Ashmead, W. H., *Mem. Carn. Mus.*, **1**, 412 (1904); Schmiedeknecht, O., *Gen. Ins.*, **97**, 37 (1909).

[21] Cresson, E. T., *Proc. Ent. Soc. Phil.*, **4**, 100 (1865); Dalla-Torre, K. W. von, *Cat. Hymen.*, **5**, 383 (1898); Schmiedeknecht, O., *Gen. Ins.*, **97**, 37 (1909).

# GROUPS IN WHICH THE SQUARES OF THE OPERATORS GENERATE A CYCLIC GROUP

BY G. A. MILLER

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS

Communicated March 23, 1936

When the squares of the operators of a given group $G$ generate a subgroup of order 2 then it is known that $G$ belongs to one of three infinite categories of groups of a very elementary structure, but when they generate a cyclic subgroup of order $2^\alpha$, $\alpha > 1$, various cases present themselves which have not yet been fully investigated.[1] It is obvious that the order of such a $G$ is of the form $2^m$, and that the number of these groups in which the operators of this cyclic subgroup are invariant under $G$ is $m - \alpha$. One and only one of these groups is abelian and its type is $\alpha + 1, 1^n$. The commutator subgroup of each of the other groups of this system is of order 2, and when $m - \alpha$ is even half of these groups involve an invariant operator of order $2^{\alpha+1}$ while no such operator appears in the rest of the groups. When $m - \alpha$ is odd then the number of these groups which involve an invariant operator of order $2^{\alpha+1}$ is $(m - \alpha + 1)/2$. A necessary and sufficient condition that such a group is abelian is that it involves no non-invariant operator of order 2.

It remains to consider these groups in which the operators of order $2^\alpha$ which are squares are not invariant and hence they are also transformed into their inverses under $G$. Such a $G$ contains a subgroup of index 2 in which these operators are invariant and which is generated by the operators of order $2^{\alpha+1}$ contained in it. This subgroup $H$ is therefore characteristic and involves all the operators whose orders exceed 4 contained in such a $G$. When $H$ is abelian there are obviously two $G$'s in which every operator of $H$ is transformed into its inverse and there is one other $G$ in which these operators are transformed into their $2^\alpha - 1$ powers. In each of these three groups the central is composed of the identity and the operators of order 2 contained in $H$. In the fourth possible non-abelian group the central is of type 2, $1^n$ and some of the operators of highest order in $H$ are transformed into their inverses while the rest are transformed into their $2^\alpha - 1$ powers under $G$. The commutator subgroup of each of these groups, as well as of all those which follow, is contained in the cyclic group of order $2^\alpha$ generated by the squares of the operators of $G$.

The number of the cyclic subgroups of order $2^{\alpha+1}$ contained in $G$ is $2^{m-\alpha-1}$ when everthe squares of the operators of $G$ are invariant under $G$ and $2^{m-\alpha-2}$ whenever this is not the case. Each of these subgroups is invariant under $G$. In fact, every cyclic subgroup of $G$ whose order exceeds 4 is invariant and when the squares of the operators of $G$ are invariant under

$G$ then every cyclic subgroup of order 4 in $G$ is also invariant under $G$. The operators of order 4 contained in $H$ generate a subgroup of index $2^{\alpha-1}$ under $H$ which includes all the operators of order 2 contained in $H$ and involves no operator whose order exceeds 4 and hence it is a characteristic subgroup of $G$. This subgroup includes the central of $G$ when $H$ is not equal to $G$ and it is generated by its operators of order 2 whenever the order of the group of inner isomorphisms of $H$ exceeds 4 and also when $H$ involves an invariant operator of order $2^{\alpha+1}$. Whenever $H$ is non-abelian the operators of $G$ which do not appear in $H$ include operators of each of the two orders 2 and 4.

Each of the operators of $G$ which is not also in $H$ transforms into themselves the operators of a subgroup of index 2 contained in the given subgroup generated by the operators of order 4 found in $H$. This subgroup of index 2 involves two and only two of the operators which are squares under $G$ but it is not restricted otherwise. It therefore has the same group of inner isomorphisms as $H$ has whenever it involves an invariant operator of order $2^{\alpha+1}$ otherwise the order of its group of inner isomorphisms is one-fourth of the order of the group of inner isomorphisms of $H$. It always has the same group of inner isomorphisms as the subgroup generated by the operators of order 4 contained in $H$, since the latter subgroup contains invariant operators of order 4 which are not found in the former, viz., the operators of order 4 which are squares under $G$ and hence are transformed into their inverses by the operators of $G$ which are not also in $H$.

When $H$ involves only one invariant operator of order 2 its group of inner isomorphisms is $2^{m-\alpha}$ whenever $m-\alpha$ is even and of order $2^{m-\alpha-1}$ when $m-\alpha$ is odd. In the latter case the group of inner isomorphisms of the subgroup of $H$ generated by its operators of order 4 is the same as the group of inner isomorphisms of $H$ while in the former case the order of the group of inner isomorphisms of this subgroup is one-fourth of that of $H$. In both of these two cases each operator of order 2 in the group of inner isomorphisms of the given subgroup corresponds to a subgroup of index 2 composed of the operators which are invariant under this isomorphism. These subgroups have a 1,1 correspondence with the subgroups of index 2, under the subgroup generated by the operators of order 4 in $H$, which involve the operators of order 4 which are squares. The number of the subgroups of index 2 in this subgroup which do not involve operators of order 4 which are squares under $G$ is therefore equal to the order of the group of inner isomorphisms of $H$ whenever $m-\alpha$ is odd. When $m-\alpha$ is even this number is equal to one-half of the order of this group of inner isomorphisms.

From the preceding paragraph it results that whenever $H$ involves only two invariant operators of order 2 then its subgroups which give rise to a cyclic quotient group of order $2^{\alpha}$ are equivalent in the sense that if such

a subgroup is used to construct a $G$ by extending $H$ by an operator which transforms its squares into their inverses, the given subgroup being composed of the operators of $H$ which correspond to themselves under this automorphism, then each of the other such subgroups is composed of all the operators of $H$ which correspond to themselves under another isomorphism of the same group. Hence we need to consider only one such automorphism for the purpose of finding all the possible $G$'s. There are therefore two and only two $G$'s for the same $H$ when $H$ involves only one invariant operator of order 2 since the isomorphism between $H$ and the commutator subgroup can be established in two and only two ways after the invariant operators under this isomorphism have been selected, and the number of the operators of orders 2 and 4 in $G$ are different for these two isomorphisms. The fact that this isomorphism can be established in only two ways results from the well known theorem that the commutators resulting from an isomorphism of order 2 correspond to their inverses thereunder.

To consider the case when $H$ is the direct product of a group $H_0$ which contains only one invariant operator of order 2 and an abelian group of type $1^n$ it is desirable to observe that every subgroup of index 2 of such a direct product is simply isomorphic either with the direct product of a subgroup of index 2 of $H$ and the given abelian subgroup of type $1^n$ or with the direct product of $H$ and an abelian group of type $1^{n-1}$. The subgroup generated by the operators of order 4 of such an $H$ again contains a subgroup of index 2 which gives rise to a cyclic quotient group of order $2^\alpha$. The latter subgroup cannot contain an operator of order 4 which is a square under $G$. When it is the direct product of such a subgroup of $H$ and an abelian group of type $1^n$ it can be selected in essentially only one way and hence there are two such $G$'s whenever $n > 0$ for the same reason as in the case noted above.

It remains to consider the case when the subgroup of $H$ which is a constituent of the given subgroup of $H$, gives rise to a cyclic quotient group of order $2^{\alpha-1}$. In this case the operators of order 4 which appear in the former subgroup and are squares are multiplied by an operator of order 2 in the given factor group of type $1^n$. As the subgroup of the latter which corresponds to the identity of the former is of order $2^{n-1}$, there is again essentially only one case that requires consideration and this gives rise to two $G$'s. It therefore results that when $H$ is a direct product and one of the factor groups is of type $1^n$, $n > 0$, there are four and only four possible $G$'s while when $n = 0$ there are only two $G$'s. It was noted above that there is one and only one group of order $2^{m-1}$, $m - 1 > \alpha$ which satisfies the three conditions that the squares of its operators generate a cyclic group of order $2^\alpha$, that it is not the direct product of such a group and an abelian group of type $1^n$, and that its squares are invariant under it. This therefore gives rise to two $G$'s. Each of the remaining $m - \alpha - 2$ groups which

satisfy the first and the last of these three conditions gives rise to four groups. Hence it results that the number of the groups of order $2^m$ which satisfy the two conditions that the squares of their operators generate the cyclic group of order $2^\alpha$ and that they transform the operators of this cyclic group into their inverses is $4(m - \alpha) - 6$. As there are $m - \alpha$ groups of order $2^m$ whose squares generate the cyclic group of order $2^\alpha$ and transform the operators of this cyclic group into themselves, it results that *there are* $5(m - \alpha) - 6$ *groups of order* $2^m$, $m > \alpha$, *which satisfy the condition that the square of their operators generate the cyclic group of order* $2^\alpha$.

[1] G. A. Miller, These PROCEEDINGS, **20**, 372–375 (1934).

---

## *MATHEMATICAL POSTULATES FOR THE LOGICAL OPERATIONS OF ASSERTION, CONJUNCTION, NEGATION AND EQUALITY*

BY EDWARD V. HUNTINGTON

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY

Communicated April 14, 1936

The purpose of this brief note is to present an "abstract deductive theory" so constructed that one of its concrete interpretations shall be ordinary logic. Details of the proofs (together with bibliographical references and acknowledgment of specific indebtedness to C. I. Lewis, W. T. Parry, A. A. Bennett and J. B. Rosser) will be given in a longer paper to be presented to the American Mathematical Society in September.

*The Abstract System.*—The "undefined concepts" of the abstract theory are $(K, T, \times, ', \equiv)$, where $K$ is a class of elements $a, b, c, \ldots$; $T$ is a subclass in $K$; $\times$ (read "times") and $\equiv$ (read "quad") are binary operators; and $'$ (read "dash") is a unary operator. (For brevity, $ab$ will be used for $a \times b$.)

The "primitive propositions" of the theory are the postulates stated below, namely, the descriptive postulates (A-E), the procedural postulates (1–7) and the formal postulates (11–14 and, if desired, 15–18).

The "theorems" of the theory follow from the postulates by purely abstract reasoning, so that any concrete system $(K, T, \times, ', \equiv)$ which has the properties stated in the postulates will have also all the properties stated in the theorems.

The "ordinary language" used in the development of the theory involves only the ideas associated with "$\longrightarrow$," "&," and "$\exists$" and the principle of substitution (see below).

*First Concrete Interpretation.*—Before proceeding to describe the abstract

theory itself, we first describe the concrete interpretation which we have chiefly in view. In this concrete system, the symbols $(K, T, \times, ', \equiv)$ are interpreted as follows.

$K$ is the class of "propositions" (better described as "sentences").

$T$ is the subclass of propositions which are "asserted," so that the statement "$a$ is in $T$" means "$a$ is asserted" (or, roughly, "$a$ is true").

$ab$ is the proposition known as the "logical product" (or "conjunction") of $a$ and $b$. The statement "$ab$ is in $T$" means "the joint proposition, $a$ and $b$, is asserted."

$a'$ is the proposition known as the "contradictory" of $a$. The statement "$a'$ is in $T$" means "the contradictory of $a$ is asserted."

$a \equiv b$ is the proposition known as the "effective equality" of $a$ and $b$, for some specified purpose. The statement "$a \equiv b$ is in $T$" means "$a$ is effectively equal to $b$," for the purpose in hand.

It should be noted that the question whether this (or any) concrete system does actually satisfy the postulates below is a question of fact, which can be decided only by observation. In this particular case, the required observations involve a critical examination of the actual usages of language, and logicians may not all agree as to the linguistic facts. The virtue of the abstract formulation of the system is that it makes this factual question precise and definite.

*Second Concrete Interpretation.*—In order to show that the postulates are at any rate *consistent*, we mention another concrete system in which the symbols $(K, T, \times, ', \equiv)$ are interpreted as follows:

$K$ = a class of four elements, 1, 2, 3, 4.
$T$ = the subclass 1, 2.
$ab$, $a'$, $a \equiv b$ = the elements shown in the adjoining tables.

| $ab$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 1 | 2 | 3 | 4 |
| 2 | 2 | 2 | 4 | 4 |
| 3 | 3 | 4 | 3 | 4 |
| 4 | 4 | 4 | 4 | 4 |

| $a$ | $a'$ |
|---|---|
| 1 | 4 |
| 2 | 3 |
| 3 | 2 |
| 4 | 1 |

| $a \equiv b$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 2 | 4 | 4 | 4 |
| 2 | 4 | 2 | 4 | 4 |
| 3 | 4 | 4 | 2 | 4 |
| 4 | 4 | 4 | 4 | 2 |

$a \equiv b$ in $T$

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | . | | | |
| 2 | | . | | |
| 3 | | | . | |
| 4 | | | | . |

(For reference, we add a "relation-table" showing when $a \equiv b$ is in $T$.)

In this case, a finite number of indubitable observations will show that all the postulates do in fact hold ("are satisfied") in this system. Hence the postulates cannot contain any mutually contradictory demands.

*Notation.*—For brevity, we shall use the notation $X \longrightarrow Y$ (read "$X$ leads to $Y$") to mean: "If, in the orderly development of the theory, we find the statement $X$ already established, then we are authorized to write down the

statement $Y$." Thus the arrow, $\longrightarrow$, is merely an abbreviation of language, and not one of the undefined symbols of the system. Obviously, if $X \longrightarrow Y$ and $Y \longrightarrow Z$, then $X \longrightarrow Z$.

It should be especially noted that the $\longrightarrow$ can stand only between statements, and not between elements of $K$. If both $X \longrightarrow Y$ and $Y \longrightarrow X$, we write $X \rightleftarrows Y$.

The notation "$X$ & $Y$" means simply "$X$ and $Y$." The notation "E$a$:(. . .)" means "there exists at least one element $a$ such that (. . .)."

*Principle of Substitution.*—It is understood that the letters $a, b, c, \ldots$ are variables, each of which may stand for any elements of $K$. That is, in any postulate or theorem, the letter $a$ may be "replaced throughout" by any expression which stands for an element of $K$.

*Descriptive Postulates.—Base:* $(K, T, \times, ', \equiv)$.

POSTULATE A. $a$ in $K$. &. $b$ in $K$: $\longrightarrow$ : $ab$ in $K$. (Read: $a$ times $b$.)
POSTULATE B. $a$ in $K$. $\longrightarrow$ . $a'$ in $K$. (Read: $a$ dash.)
POSTULATE C. $a$ in $T$. $\longrightarrow$ . $a$ in $K$.
POSTULATE D. $\exists a$: ($a$ in $T$).
POSTULATE E. $a$ in $K$. &. $b$ in $K$: $\longrightarrow$: $a \equiv b$ in K. (Read: $a$ quad $b$.)
DEFINITION 1. $a = b$ . $\rightleftarrows$ . $a \equiv b$ in $T$.

It should be noted that the symbol "=" introduced by Definition 1 is part of the abstract system and has no properties except such as are deducible from the postulates governing the symbols $\equiv$ and $T$. The notation $a = b$ is merely a convenient abbreviation for the statement $a \equiv b$ in $T$.

*Procedural postulates*

POSTULATE 1. $a$ in $T$. &. $b$ in $T$ : $\longrightarrow$ : $ab$ in $T$. (Rule of adjunction.)
POSTULATE 2. $ab$ in $T$ : $\longrightarrow$ : $a$ in $T$. (Rule of simplification for $\times$.)
POSTULATE 3.[1] $a = b$. & . $b = c$: $\longrightarrow$ : $a = c$. (Rule of transitivity for =.)
POSTULATE 4. $a$ in $T$ . &. $a = b$ :$\longrightarrow$: $b$ in $T$.
POSTULATE 5. $a = b$ :$\longrightarrow$: $ac = bc$.
POSTULATE 6. $a = b$ :$\longrightarrow$: $a' = b'$.
POSTULATE 7. $a = b$ :$\longrightarrow$: $(a \equiv c) = (b \equiv c)$.

*Formal postulates (Part I.)*

POSTULATE 11. $ab = ba$. (Commutative law for $\times$.)
POSTULATE 12. $(ab)c = a(bc)$. (Associative law for $\times$.)
POSTULATE 13. $(a'b)'(a'b')' = a$. (Law of expansion.)
POSTULATE 14. $(a \equiv b) = (b \equiv a)$. (Commutative law for $\equiv$.)

*Theorems Deducible from Part I.*—First of all we prove the following theorem which shows that the sign =, introduced in Definition 1, may be handled just as if it were the ordinary equality sign of mathematics.

20. *Theorem of Replacement.*—Let $\xi$ be any expression which stands for

an element of the class $K$ (that is, any expression built up out of parentheses, the letters, $a, b, c, \ldots$ and the signs $\times, ', \equiv$, in accordance with Postulates A, B, E); and suppose that at some point in the expression $\xi$ there occurs the letter $a$. Let $\eta$ be another expression which differs from $\xi$ only in having this particular $a$ replaced by a new letter $b$. Then the following procedural theorem will hold:

$$a \equiv b \text{ in } T .\longrightarrow. \xi \equiv \eta \text{ in } T, \quad \text{or,} \quad a = b .\longrightarrow. \xi = \eta,$$

where $\xi$ and $\eta$ are expressions which differ only in having some one $a$ replaced by $b$.

In the light of this theorem, Postulates A, B, 11, 12, 13 make the system a Boolean algebra with respect to $(K, \times, ')$. Hence we have:

21. $a'' = a$. (Law of double negation.)
22. $aa' = bb'$.

DEFINITION 2. $Z = aa'$. (The "zero-element" of the system.)
DEFINITION 3. $U = Z'$. (The "universe element.")

23. $aZ = Z$.
24. $aZ' = a$.
25. $aa = a$. (Idempotent law for $\times$.)
26. $a = a$. (Law of identity.)

DEFINITION 4. $(a \vee b) = (a'b')'$. (Read: $a$ wedge $b$.)

27. $a(a \vee b) = a$.
28. $a(b \vee c) = ab \vee ac$. (Distributive law for $\times$.)
29. $a \vee bc = (a \vee b)\,(a \vee c)$. (Distributive law for $\vee$.)

DEFINITION 5. $(a < b) .\rightleftarrows. ab' = Z$. (Read: $a$ less than $b$).

30. $ab' = Z .\rightleftarrows. ab = a$.

This last theorem shows that Definition 5 (which is Lewis' "strict implication") is equivalent to the following:

31. $a<b .\rightleftarrows. ab = a$. ("Effective implication.")

Hence $a<b$ has the fundamental properties commonly associated with the word "implication," as follows:

32. $(a < b)\ \&\ (b<c) .\longrightarrow. (a<c)$. (Syllogism.)
33. $(a \text{ in } T)\ \&\ (a<b) .\longrightarrow. (b \text{ in } T)$. (Inference.)
34. $(a<b) .\rightleftarrows. (b'<a')$.
35. $(a<b)\ \&\ (b<a) .\rightleftarrows. (a = b)$.

We note also:

36. $Z'$ is in $T$.
37. $Z$ in $T .\longrightarrow. a$ in $T$.

That is, if $Z$ is in $T$, then every element of $K$ is in $T$. This is a trivial case, which may be referred as "$T = K$."

38. $Z < a$.
39. $a < Z'$.

*Supplementary Definitions.*—The following definitions, which are characteristic of Lewis' theory, will be useful in formulating the postulates of Part II.

DEFINITION 6. $a^* . = . a \equiv Z$. (Read: $a$ star.)

By Definition 1, this gives the theorem

41. $a^*$ in $T .\rightleftarrows. a \equiv Z$ in $T .\rightleftarrows. a = Z$.

DEFINITION 7. $(a \strictif b) . = . (ab') \equiv Z$.

By Definition 6, this gives $(a \strictif b) = (ab')$,* and by Definitions 1, 5,

42. $a \strictif b$ in $T .\rightleftarrows. ab' = Z .\rightleftarrows. a < b$. Also, by 25 and 21,
43. $a \strictif a' = a^*$.
44. $a \strictif a = Z^*$.

*Formal postulates (Part II).*

POSTULATE 15. $(a \strictif b)(b \strictif a) = (a \equiv b)$.
POSTULATE 16. $(a \strictif b)(b \strictif c) \strictif (a \strictif c)$ is in $T$.
POSTULATE 17. $a(a \strictif b) \strictif b$ is in $T$.
POSTULATE 18.[1] $(a \strictif a') \strictif (a \strictif b)$ is in $T$.

*Theorems Depending on Part II.*—Of the theorems which require the postulates of Part II for their proof, the following is the most characteristic:

45. $a \strictif b . = . ab \equiv a$.

It will be noted that 31 follows from 45 immediately; but 45 cannot be deduced from 31 without the aid of Part II.

Further theorems of Part II are:

46. $a^*a = Z$.
47. $Z^{**} = Z$.
48. $(a \equiv a) = Z^*$
49. $Z^*$ is in $T$.
50. $a^* = a^*Z^*$.
51. $a \strictif Z^* = Z^*$.

52. $(a \strictif b)$ in $T .\longrightarrow. (a \strictif b) = Z^*$.
53. $(a \equiv b)$ in $T .\longrightarrow. (a \equiv b) = Z^*$.

*Concrete Interpretation of $Z$, $Z'$, $Z^*$; $a^*$, and $a \strictif b$.*—If the symbols $(K, T, \times, ', \equiv)$ are interpreted as in the first example above, then $Z$

means "absurdity" (the conjunction of a proposition, $a$, with its contradictory, $a'$); $Z'$ means "tautology" (the contradiction of absurdity); $a^*$ means the proposition known as "the absurdity of $a$"; and the statement "$a^*$ is in $T$" means: "the absurdity of $a$ is asserted," or simply, "$a$ *is* absurd." Further, $a \mapsto b$ means the proposition known as "the (strict) implication of $a$ toward $b$"; and the statement "$a \mapsto b$ is in $T$" means: "the (strict) implication of $a$ toward $b$ is asserted," or simply, "$a$ does (strictly) imply $b$." This is equivalent to the statement "$ab'$ is absurd." Finally, by 48, $Z^*$ means "certainty" (that is, the effective equality of $a$ and $a$).

If, on the other hand, the symbols ($K$, $T$, $\times$, $'$, $\equiv$) are interpreted as in the second example above, then $Z$ means the element 4; $a^*$ and $a \mapsto b$ mean the elements indicated in the adjoining tables ($a$ on the left, $b$ along the top); $Z'$ means 1; and $Z^*$ means 2. For reference, we add a "relation table" showing when $a \mapsto b$ is in $T$.

| $a$ | $a^*$ |
|---|---|
| 1 | 4 |
| 2 | 4 |
| 3 | 4 |
| 4 | 2 |

| $a \mapsto b$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 2 | 4 | 4 | 4 |
| 2 | 2 | 2 | 4 | 4 |
| 3 | 2 | 4 | 2 | 4 |
| 4 | 2 | 2 | 2 | 2 |

$a \mapsto b$ in $T$

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | . | | | |
| 2 | . | . | | |
| 3 | . | | . | |
| 4 | . | . | . | . |

It may be noted that since always $Z^* \mapsto Z'$ is in $T$, but not always $Z' \mapsto Z^*$ in $T$, it comes out that "certainty" ($Z^*$) is logically weaker than "tautology" ($Z'$).

It would be interesting to investigate in detail how much of practical value is added to Lewis' system by the inclusion of this supplementary part, which makes possible a sharp distinction between $a'$ (the contradictory of $a$) and $a^*$ (the absurdity of $a$). [In the Whitehead-Russell theory of material implication, where $a \supset b. = . (ab')'$, this distinction is not available.] In this connection it should be noted that the so-called paradoxes expressed in the formulas

P1. $a$ in $T$. & .$b$ in $T :\longrightarrow : a = b$; P3. $a$ in $T .\longrightarrow . (b \mapsto a)$ in $T$;
P2. $(a \mapsto b)'$ in $T .\longrightarrow . (b \mapsto a)$ in $T$; P4. $a'$ in $T.\longrightarrow .(a \mapsto b)$ in $T$;

are all avoided in Lewis's theory, since none of these formulas can be deduced from the postulates given above.

[1] Postulates 3 and 18 prove to be redundant.

# ISOMETRIES OF 2-DIMENSIONAL RIEMANNIAN MANIFOLDS INTO THEMSELVES

BY SUMNER BYRON MYERS*

PRINCETON UNIVERSITY AND THE INSTITUTE FOR ADVANCED STUDY

Communicated March 26, 1936

1. *Introduction.*—This abstract contains the results of a study of the isometries of 2-dimensional Riemannian manifolds[1] into themselves. Such a study is essentially a combination of a study of groups of motions of Riemannian spaces in the small[2] with a study of topological transformations of 2-dimensional manifolds into themselves.[3] The subject comes into close contact with the theory of conformal mapping[4] and the theory of isometries of general metric spaces into themselves.[5] We show here that the group of isometries of a compact 2-dimensional Riemannian manifold is either continuous (i.e., locally euclidean) or finite; the continuous groups are studied by means of methods initiated by W. Rinow,[6] while a method is devised for the study of periodic isometries by means of the "minimum point locus" studied by the author in previous papers.[7]

2. *General Notions about Isometries of n-Dimensional Riemannian Manifolds.*—Let $M$ be a regular $n$-dimensional manifold of class 5 with a line element of class $C^4$. An isometry of $M$ into itself is defined to be a length-preserving homeomorphism. We make $M$ into a metric space by defining the distance between two points to be the greatest lower bound of the lengths of rectifiable arcs joining the two points.

THEOREM 1. *The distance preserving homeomorphisms of an n-dimensional Riemannian manifold $M$ are identical with the length-preserving homeomorphisms.*

The isometries of $M$ into itself form a metrizable group $G$. If $M$ is compact, the distance between two isometries $T_1$ and $T_2$ can be defined to be the maximum distance $\overline{T_1(P)\cdot T_2(P)}$ for $P$ on $M$, and the group is a metric[8] group.

THEOREM 2. *A discrete set of isometries of an n-dimensional Riemannian manifold into itself carries any point into a discrete set of points. Hence if the manifold is compact its group of isometries is compact.*[8]

The universal covering manifold $\overline{M}$ of $M$ becomes a Riemannian manifold if each neighborhood in $\overline{M}$ is provided with the same Riemannian metric as the neighborhood in $M$ which it covers.

THEOREM 3. *Let $G$ be the group of isometries of $M$ into itself, $\overline{G}$ the group of isometries of $\overline{M}$ into itself. Let $H$ be the discrete group of fixed point free isometries of $\overline{M}$ into itself isomorphic to the fundamental group of $M$ such that identification of points congruent under $H$ produces $M$. Then $G = K/H$,*

*where $K$ is the largest subgroup of $\bar{G}$ which contains $H$ as an invariant subgroup.*[9]

3. *2-Dimensional Riemannian Manifolds.*—An isometry $T$ of a 2-dimensional Riemannian manifold into itself carries geodesics into geodesics and preserves angle. If $O$ is a fixed point of $T$, then the behavior under $T$ of the geodesic rays issuing from $O$ can be described either by $\phi = \phi' +$ constant of $\phi = -\phi'$. In the first case, we call $T$ a rotation about $O$; in the second case, we call $T$ a reflection. A reflection is an involution, and reverses the indicatrix on an orientable surface. Under a reflection there is a pointwise fixed geodesic, which can be shown to be a closed geodesic if the manifold is compact. On a non-orientable surface an isometry may be at once a reflection and a rotation.

THEOREM 4. *The group of isometries of a 2-dimensional Riemannian manifold is either discrete or locally euclidean.*

THEOREM 5. *The group of isometries of a 2-dimensional Riemannian manifold is discrete, except possibly in the case of surfaces homeomorphic to the sphere, projective plane, plane, cylinder, non-orientable cylinder (Mobius strip), torus and Klein bottle.*

From Theorems 5 and 2 follows:

THEOREM 6. *The group of isometries of a compact 2-dimensional Riemannian manifold of negative Euler characteristic is finite, and hence every isometry is periodic.*

The exceptional surfaces named in Theorem 5 can all have continuous Lie groups of isometries as subgroups of their complete isometry groups, if provided with the proper Riemannian metric. These groups can be at most 3-parameter, and can be 2- or 3-parameter groups only if the surface is complete and of constant curvature. Conditions on the metric which are necessary for the existence of a 1-parameter group are also well known.[10] The sphere (of constant positive curvature)[11] has a transitive 3-parameter group of rotations with two fixed points, the ellipsoid of revolution a 1-parameter group of such rotations. The manifolds (projective planes) obtained from these two by identification of diametrically opposite points have respectively 3- and 1-parameter groups of rotations with one fixed point; the rotations through 180° are also reflections in a closed geodesic. The euclidean plane and hyperbolic plane have well-known transitive 3-parameter groups of isometries, while a surface of non-constant curvature homeomorphic to the plane may have either a 1-parameter group of rotations with one fixed point or a 1-parameter group of translations. The ordinary cylinder of revolution has a transitive 2-parameter group of translations, while the hyperbolic cylinder (of constant negative curvature) has a 1-parameter group of fixed point free isometries. A surface of non-constant curvature homeomorphic to the cylinder may also have a 1-parameter group of fixed point free isometries. The non-orientable cylinder

(unbounded Möbius strip) of zero curvature has a 1-parameter fixed point free group, as does the homeomorphic surface of constant negative curvature. A non-orientable cylinder of non-constant curvature may also have a 1-parameter group of fixed point free isometries. The torus of zero curvature has a 2-parameter group of translations; a torus of non-constant curvature may have a 1-parameter group. A Klein bottle of zero curvature has a 1-parameter group of translations, and a Klein bottle of non-constant curvature may also have a 1-parameter group.

THEOREM 7. *Every isometry of a 2-dimensional Riemannian manifold is topologically (in fact, conformally) equivalent to some isometry of a homeomorphic surface of constant curvature.*

THEOREM 8. *If $T$ is an isometry of a 2-dimensional Riemannian manifold with a fixed point $O$, the minimum point locus with respect to $O$ goes into itself under $T$.*

THEOREM 9. *If $T$ is an isometry of a 2-dimensional Riemannian manifold with a fixed point $O$, every other fixed point of $T$ lies on the minimum point locus $m$ with respect to $O$, unless $T$ is a reflection in a geodesic through $O$, in which case all other fixed points except the points on $g$ lie on $m$.*

According to previous results of the author,[12] on a compact surface of constant curvature the minimum point locus with respect to any point $O$ consists of a linear graph which is a base for non-bounding 1-cycles. Theorems 7, 8 and 9 tell us that the study of isometries of a compact surface $S$ with a fixed point $O$ can be reduced to the study of transformations into itself of a linear graph which is a base for non-bounding 1-cycles on $S$. This idea will be exploited in a later paper, and applied to a study of general periodic topological transformations.

* NATIONAL RESEARCH FELLOW.

[1] For definition, see Hopf and Rinow, "Ueber den Begriff der vollständigen differentialgeometrische Fläche," *Commentarii Mathematici Helvetici*, 3 (1931), p. 209, and also Veblen and Whitehead, "The Foundations of Differential Geometry."

[2] Eisenhart, *Riemannian Geometry*, Chapter 6.

[3] Kerekjarto, *Vorlesungen über Topologie*, Chapter 6.

[4] See Weyl, *Die Idee der Riemannschen Fläche*, §21. The connection is due to the fact that every orientable 2-dimensional Riemannian manifold is a Riemann surface whose conformal transformations are the angle and sense preserving homeomorphisms of the Riemannian manifold and hence include the isometries.

[5] See van Dantzig and van der Waerden, "Über metrisch homogene Raüme," *Hamburg Univ. Math. Sem. Abhandl.*, **6**, 367–376 (1928).

[6] W. Rinow, "Ueber Flächen mit Verschiebungselementen," *Math. Ann.*, **107**, 95–112 (1932).

[7] S. B. Myers, "Connections between Differential Geometry and Topology. I. Simply Connected Surfaces," *Duke Math. Jour.*, **1**, 376–391 (1935), and "II. Closed Surfaces," *Duke Math. Jour.*, **2** (1936).

[8] See ref. 5, p. 369.

[9] This theorem is true if the word "isometries" is replaced by "topological transformations." Cf. ref. 4, pp. 163–164.

[10] See ref. 2, pp. 241–244.

[11] For the classification of topological types of surfaces according to the kind of constant curvature possible on them, see Hopf, "Differentialgeometrie und topologische Gestalt," *Jahr. Deutsch. Math. Verein.*, **41**, 212–213 (1932).

[12] See ref. 7.

---

## ON THE CONNECTIVITY RING OF A BICOMPACT SPACE

By J. W. Alexander

The Institute for Advanced Study

Communicated March 30, 1936

1. *A cartesian space* $C$ will be the space of any finite or transfinite set of real variables $x_i$, topologized in the following manner. If $p^\circ$ is any point of $C$, $\epsilon$ any positive number, and $x_{i_s}$ $(s = 1, 2, \ldots, n)$ any finite set of coördinates, arbitrarily selected from the entire set, there will be a neighborhood of $p^\circ$ consisting of all points $p$ of $C$ such that their coördinates $x_{i_s}$ satisfy the inequalities

$$| x_{i_s} - x^\circ_{i_s} | < \epsilon \qquad (s = 1, 2, \ldots, n.)$$

Moreover, the set of all neighborhoods of the above type will be a basis for the open sets of $C$. The topology assigned to $C$ will thus be the so-called *weak* topology. The *dimensionality* $\alpha$ of $C$ will be equal to the number of coördinates $x_i$ of a point $p$ of $C$. It will either be a finite or a transfinite cardinal.

A point set $D$ of $C$ will be said to be *bounded* if each coördinate $x_i$ of a variable point $p$ of $D$ is bounded. Thus, the set $D$ will be bounded if, and only if, it lies within some parallelotope determined by imposing a condition of the form

$$-m_i \leqq x_i \leqq m_i \qquad (m_i \text{ constant}) \qquad (1)$$

on each of the coördinates $x_i$. It can be shown that every closed, bounded point set $B$ in $C$ is a bicompact, topological space.[1] Furthermore, every bicompact, topological space may be regarded as a closed, bounded point set $B$ immersed in a suitable cartesian space $C$.[2] For example, we may form the set of all continuous (and therefore bounded) functions $x_i$ on $B$ and regard the values of the $x_i$'s at a point $p$ of $B$ as the coördinates of $p$ in the desired space $C$.

2. Now, let $C$ be any cartesian space of dimensionality $\alpha$. Then if we take any finite set of distinct linear functions of the form

$$x_i - c_{ij} \qquad (c_{ij} \text{ constant}) \qquad (2)$$

these functions will determine a *rectangular grating* $\Gamma$ of $C$ composed of certain point sets, called (generalized) *cells*, which will be defined as follows. With each of the functions $x_i - c_{ij}$ we shall associate a function $\sigma_{ij}(p)$ taking on the values 1, 0 or $-1$ at a point $p$ according as $x_i - c_{ij}$ is positive, zero or negative at $p$. The determinations of the various functions $\sigma_{ij}$ at $p$ will be called the *signature* of $p$. Each point $p$ of $C$ will belong to a (generalized) *cell* $E$ of the grating consisting of all points with the same signature as $p$. The cell $E$ will be called an $(\alpha - k)$-cell if the number of zeros in the signature of $p$ is precisely $k$. A cell $E$ will be said to be *on the boundary* of a cell $F$ if the signature of a point of $E$ is obtainable from the signature of a point of $F$ by changing the values of one or more of the functions $\sigma_{ij}$ from $\pm 1$ to 0. If the dimensionality $\alpha$ of $C$ is finite, the grating $\Gamma$ will be a complex of the ordinary, classical type, and an $(\alpha - k)$-cell of $\Gamma$ will be an ordinary cell (possibly unbounded) of dimensionality $\alpha - k$. If the dimensionality of $C$ is transfinite, the grating $\Gamma$ will be an aggregate with all the essential formal properties of a complex, but composed of generalized cells of symbolic dimensionalities $\alpha - k$. In either case, the formal theory of chains and cycles will be applicable to $\Gamma$, and we shall be able to speak of $(\alpha - k)$-chains, $(\alpha - k)$-cycles, homologies and so on. If we make a refinement $\Gamma'$ of the grating $\Gamma$ by adjoining a finite number of new linear functions to the set (2), we shall be able to identify each chain $K$ of $\Gamma$ with a refinement $K'$ of $K$ in $\Gamma'$.

The theory of intersecting chains may also be extended, at once, to gratings. If we wish to form the *intersection* $KL$ of an $(\alpha - k)$-chain $K$ with an $(\alpha - l)$-chain $L$ we can use the following geometrical procedure. First we construct a "displacement" $\Gamma'$ of $\Gamma$ determined by changing the functions (2) to

$$x_i - c_{ij} - \epsilon \qquad (\epsilon \text{ positive and sufficiently small}) \qquad (3)$$

then we replace the chain $L$ by its displacement $L'$ in $\Gamma'$. The chains $K$ and $L'$ are now in general positions with respect to one another and therefore intersect in an $(\alpha - k - l)$-chain $KL'$ belonging to the grating determined by the combined functions (2) and (3). (Here we must make the customary conventions about the sign of the intersection.) The desired intersection $KL$ will now be the $(\alpha - k - l)$-chain of $\Gamma$ which is approached by $KL'$ as the parameter $\epsilon$ of the displacement approaches 0. The chain $KL$ may, of course, also be defined directly in purely combinatory terms, without the intervention of the auxiliary grating $\Gamma'$.

3. Now, let $A$ be any open set of points in a cartesian space $C$. Then a chain $K$ of any grating $\Gamma$ of $C$ will be said to *belong to* $A$ if the points and limit points of the underlying cells of $K$ belong to $A$. Moreover, such a chain will be called a *cycle of* $A$ if it is a cycle of $\Gamma$. A cycle $K$ of $A$ will be said to be *homologous to zero*, $K \sim 0$, in $A$ if it bounds a chain $L$

of $A$ on any refinement $\Gamma'$ of $\Gamma$. (Here, of course, we are identifying the chain $K$ with its refinement $K'$ in $\Gamma'$.) Two chains $K$ and $K'$ belonging to the same or to different gratings $\Gamma$ and $\Gamma'$, respectively, will be said to be homologous, $K \sim K'$, if their difference is homologous to zero in any common refinement of $\Gamma$ and $\Gamma'$. The *$(\alpha - k)$-th connectivity group* of $A$ will be the group of all cycles of $A$ modulo the subgroup of all cycles of $A$ that are homologous to zero in $A$.

THEOREM. *The intersection $KL$ of two cycles $K$ and $L$ of $A$ is always homologous to zero, $KL \sim 0$.*

For there exists a chain $M$ of $C$, though not necessarily of $A$, such that $M$ is bounded by $K$,

$$\beta M = K \qquad \text{(in } C\text{)}.$$

Moreover, by a well-known theorem on intersections,

$$\beta(ML) = (\beta M)L = KL \qquad \text{(in } C\text{)}.$$

Finally, since $L$ belongs to $A$, the intersection $ML$ must also belong to $A$. Therefore,

$$\beta(ML) = KL \sim 0 \qquad \text{(in } A\text{)}.$$

Thus, for open sets in cartesian space the ordinary intersection theory is trivial.

4. Now, let us take any bicompact, topological space $B$, which we picture as a closed, bounded point set in a cartesian space $C$, or rather in the space $C^*$ obtained from $C$ by removing any one of its points. (We work in the space $C^*$ rather than in $C$, because $C^*$ has the property that its connectivity groups all reduce to the identity. Cf. below.) Let $A$ be the open complement of $B$ with respect to $C^*$. Then it can be shown that the $(\alpha - k)$-th connectivity group of $A$ is a topological invariant of $B$. We shall call this invariant the *$k$-th connectivity group* of $B$. (The $k$-th Vietoris group of $B$, for cycles with real coefficients, reduced modulo 1, is the character group of the $k$-th connectivity group as here defined.)

It is also possible to define the *product* of two group elements $\varphi$ and $\psi$ belonging to the same or to different connectivity groups of $B$ and thus to obtain what is essentially a *connectivity ring* of $B$. Let $K$ and $L$ be any two cycles of $A$ determining the group elements $\varphi$ and $\psi$, respectively, and let $K^*$ and $L^*$ be chains of $C^*$ bounded by $K$ and $L$, respectively. Then we may define the product $\varphi\psi$ as the group element determined by the boundary $\beta(K^*L^*)$ of the intersections of $K^*$ with $L^*$. The cycle $\beta(K^*L^*)$ obviously belongs to $A$. For by a well-known theorem on intersections,

$$\begin{aligned}\beta(K^*L^*) &= (\beta K^*)L^* \pm K^*(\beta L^*) \\ &= KL^* \pm K^*L.\end{aligned}$$

Moreover, this last chain obviously belongs to $A$, since $K$ and $L$ both belong to $A$. It is easy to show that the element $\varphi\psi$ is independent of the choice of the cycles $K$ and $L$ determining $\varphi$ and $\psi$, respectively, and of the chains $K^*$ and $L^*$ bounded by $K$ and $L$, respectively. However, the proof depends essentially on the fact that the connectivity groups of $C^*$ all reduce to the identity. This is why we work in $C^*$ and not in $C$. It will be oberved that if we had defined the group element $\varphi\psi$ merely as the element determined by the intersections of $K$ with $L$ we would have obtained nothing of interest. For, by §3, we always have $KL \sim 0$.

The proof of the invariance of the connectivity groups and connectivity ring of $B$ is quite simple. We compare two different replicas $B_1$ and $B_2$ of $B$ in two different cartesian spaces $C_1$ and $C_2$ with the replica $B_0$ determined by $B_1$ and $B_2$ in the product space $C_0 = C_1C_2$. It is easy to show that the invariants determined by the replica $B_0$ are the same as the ones determined by its projections $B_1$ and $B_2$. The details of the proof will be given elsewhere.

[1] By a topological space we here mean a space satisfying the first four axioms of Hausdorff, though not necessarily the countability axioms. Every bicompact topological space is a normal space.

[2] Compare, for example, A. Tychonoff: "Über die Topologische Erweiterung von Räumen," *Math. Ann.*, **102**, 544–551 (1929–30).

---

## *CONFORMAL GEOMETRY OF HORN ANGLES*[1]

BY EDWARD KASNER AND GEORGE COMENETZ

DEPARTMENT OF MATHEMATICS, COLUMBIA UNIVERSITY

Communicated April 14, 1936

The object of this paper is to derive a fundamental identity in conformal plane geometry, relating the conformal invariants of adjacent horn angles. The identity is expressed in two different forms by equations (16) and (18) below. A horn angle has a unique conformal invariant of third order. Three curves give rise to a fundamental inequality (20), and four curves to an identity relation.

1. A horn angle (or curvilinear angle of magnitude zero), is the figure formed by an ordered pair of (analytic) curves which pass through a point in the same direction. The contact will be assumed to be of the first order, so that the curvatures $\gamma_1$ and $\gamma_2$ of the two curves at the point are distinct. Then if $d\gamma_1/ds_1$, and $d\gamma_2/ds_2$ denote the derivatives of curvature with respect to arc-length, the quantity

$$I_{12} \equiv \frac{d\gamma_2/ds_2 - d\gamma_1/ds_1}{(\gamma_2 - \gamma_1)^2} \tag{1}$$

is termed *the conformal invariant* of the horn angle. The justification for this name is that two horn angles are equivalent under the group of all conformal transformations in the plane if, and only if, they have the same value of $I_{12}$. (By a conformal transformation is to be understood a one-to-one transformation defined in some region of the plane by $w = f(z)$, where $f(z)$ is a function of a complex variable analytic throughout the region with $f'(z) \neq 0$ at all points. The figure transformed is to lie in the *interior* of the region*.) So the conformal invariant $I_{12}$ is the unique conformal invariant of a horn angle.[2] Next to the magnitude of an ordinary curvilinear angle, it is the *simplest example of a conformal invariant.* It follows of course that it is invariant in particular under the group of inversions.

2. We shall give here the proof that *I is invariant under any (regular) conformal transformation.*[3]

It is sufficient to consider horn angles formed by two curves tangent to the $x$-axis at the origin, since (1) is certainly a euclidean invariant. Let the equation of one of the curves be ,

$$y = a_2 x^2 + a_3 x^3 + \ldots \tag{2}$$

The conformal transformation will be given in some neighborhood of the origin by

$$w = c_1 z + c_2 z^2 + c_3 z^3 + \ldots \tag{3}$$

with $c_1 \neq 0$, where $z = x + iy$, $w = X + iY$, and the coefficients are complex numbers, say $c_n = \alpha_n + i\beta_n$. But $c_1$ is real, since the horizontal direction is to remain fixed. Let the transformed curve be

$$Y = A_2 X^2 + A_3 X^3 + \ldots \tag{4}$$

Separating (3) into real and imaginary parts, we obtain the equations of transformation in the real form

$$\begin{aligned} X &= \alpha_1 x + \alpha_2 x^2 - 2\beta_2 xy - \alpha_2 y^2 + \ldots, \\ Y &= \alpha_1 y + \beta_2 x^2 + 2\alpha_2 xy - \beta_2 y^2 + \beta_3 x^3 + \ldots \end{aligned} \tag{5}$$

Substituting in (4) and using (2), we derive an identity in $x$. Equating the coefficients of $x^2$ and $x^3$, respectively, we find

$$A_2 = (\alpha_1 a_2 + \beta_2)/\alpha_1^2, \; A_3 = (\alpha_1^2 a_3 - 2\alpha_2 \beta_2 + \alpha_1 \beta_3)/\alpha_1^4. \tag{6}$$

Hence, employing primed letters for the other curve of the horn angle, we have

$$A_2 - A_2' = (a_2 - a_2')/\alpha_1, \; A_3 - A_3' = (a_3 - a_3')/\alpha_1^2. \tag{7}$$

Since the curves are horizontal at the origin, the curvature $\gamma$ there is $2a_2$ and $d\gamma/ds = 6\,a_3$. We can therefore replace (7) by

$$\Gamma - \Gamma' = (\gamma - \gamma')/\alpha_1, \; d\Gamma/dS - d\Gamma'/dS' = (d\gamma/ds - d\gamma'/ds')/\alpha_1^2, \tag{8}$$

where the capital letters are used for the transformed curves. It follows at once that the quotient in (1) has the same value in the capital letters as in the small letters, and this completes the proof.

THEOREM. *The relations* (8) *mean that* $\gamma - \gamma'$ *and* $d\gamma/ds - d\gamma'/ds'$ *are relative conformal invariants* (the factor $\alpha_1^2$ is obviously the Jacobian of (5)).

The relative invariance of $\gamma - \gamma'$ has recently been proved independently by Ostrowski, who gives several very interesting applications.[4] Ostrowski's expression of second order, and the other expression of third order are in fact the only relative invariants for the general case of contact of first order.

Another expression for the absolute conformal invariant of a horn angle can be given in terms of more directly measurable quantities than those appearing in (1). If $\rho_1$ and $\rho_2$ denote the radii of curvature of the sides of the angle, and $\delta_1$ and $\delta_2$ denote the angles of deviation, then the expression is

$$I_{12} = 3\,(\rho_2^2 \tan \delta_1 - \rho_1^2 \tan \delta_2)/(\rho_2 - \rho_1)^2. \tag{1'}$$

This is derived from (1) by use of the definition $\rho = 1/\gamma$ and the identity $\tan \delta = {}^1/_3\, d\rho/ds$, which is easy to prove from the definition of deviation.[5]

3. If three curves pass through a point in different directions, they form three curvilinear angles of magnitudes $\theta_{12}$, $\theta_{23}$ and $\theta_{13}$ ($\theta_{ij}$ is the angle from the $i^{th}$ to the $j^{th}$ curve). These three quantities are not independent, for the largest angle equals the sum of the other two, or more symmetrically we have (modulo $2\pi$ or 360°)

$$\theta_{12} + \theta_{23} + \theta_{31} = 0. \tag{9}$$

Our problem now is this: to find the identity which corresponds to (9) when we consider horn angles instead of ordinary angles, and use the conformal invariant instead of radian or degree measure.

Suppose therefore that three curves are tangent at a point so as to form three horn angles whose conformal invariants are $I_{12}$, $I_{23}$ and $I_{31}$. By the definition (1) we have (writing $\eta$ for $d\gamma/ds$)

$$I_{12} = (\eta_2 - \eta_1)/(\gamma_2 - \gamma_1)^2, \; I_{23} = (\eta_3 - \eta_2)/(\gamma_3 - \gamma_2)^2, \; I_{31} = (\eta_1 - \eta_3)/(\gamma_1 - \gamma_3)^2. \tag{10}$$

Now to find an identity among the three $I$'s means, in effect, to eliminate the six independent letters $\gamma_i$, $\eta_i$ from the three equations (10). But it is easy to show that this cannot be done. Hence *for the case of three curves there is no identity like* (9). However, there is an *inequality* for three curves, stated at the end of the paper.

The next step therefore is to draw a fourth curve tangent at the same point to the other three, forming three new horn angles. This introduces two more letters, $\gamma_4$ and $\eta_4$, but three more equations

$$I_{34} = (\eta_4 - \eta_3)/(\gamma_4 - \gamma_3)^2,\ I_{14} = (\eta_4 - \eta_1)/(\gamma_4 - \gamma_1)^2,\ I_{24} = (\eta_4 - \eta_2)/(\gamma_4 - \gamma_2)^2. \quad (11)$$

It turns out that a unique eliminant now exists.

To prove this, we define six new letters $x_i$, $y_i$ by

$$x_i = \gamma_4 - \gamma_i,\ y_i = \eta_4 - \eta_i \quad (i = 1, 2, 3). \quad (12)$$

Also, for convenience we make the following abbreviations:

$$I_{12} = a,\ I_{23} = b,\ I_{31} = c,\ I_{34} = d,\ I_{14} = e,\ I_{24} = f; \quad (13)$$

$$g = a - e,\ h = a + f,\ j = b - f,\ k = b + d,\ l = c - d,\ m = c + e \quad (14)$$

By elimination of the $y$'s, (10) and (11) are readily reduced to

$$\begin{aligned} g\,x_1^2 - 2\,a\,x_1 x_2 + h\,x_2^2 &= 0, \\ j\,x_2^2 - 2\,b\,x_2 x_3 + k\,x_3^2 &= 0, \\ l\,x_3^2 - 2\,cx_3x_1 \quad + mx_1^2 &= 0. \end{aligned} \quad (15)$$

The second and third of these equations are solved as quadratics for the ratios $x_2/x_3$ and $x_1/x_3$, respectively, and the answers are substituted into the first equation. The following eliminant is obtained after some calculation:

$$\begin{aligned} &(4a^2jklm - 4abcgjl - 4\,abchkm + 8\,b^2c^2gh - 4b^2ghlm - 4c^2ghjk + \\ &\quad g^2j^2l^2 + 2ghjklm + h^2k^2m^2)^2 \\ &\quad - 16\,(b^2 - jk)\,(c^2 - lm)\,(agjl + ahkm - 2bcgh)^2 = 0. \end{aligned} \quad (16)$$

This equation in terms of the abbreviations (13), (14) is the relation sought for. It is the identity which the six values of the conformal invariant must satisfy.

To write this identity in the explicit form analogous to (9), we substitute back the letters $I_{12}, \ldots, I_{24}$ from (13), (14) and multiply out. After considerable computation we find for the result an extremely long equation,

$$I_{12}^4\,I_{23}^4\,I_{34}^4 + \ldots - 372\,I_{12}^2\,I_{23}^2\,I_{31}^2\,I_{34}^2\,I_{14}^2\,I_{24}^2 = 0. \quad (17)$$

There are 805 terms (too many to print) in the left member. It is homogenous of the twelfth degree in the six $I$'s; and is of degree 4 in each of the $I$'s taken separately. It is not completely symmetrical in the $I$'s, but it does transform into itself under a certain group $G$ of 24 substitutions, isomorphic to the symmetric group on four letters. These substitutions are in fact obtained by noting the effect of all the permutations of (1, 2, 3, 4) on the subscripts of the six $I$'s. Such permutations amount to a renumbering of the four curves which are the sides of the six horn angles. For ex-

ample, the permutation (1, 2, 4, 3) gives rise to the substitution $I_{12} \longrightarrow I_{12}$, $I_{23} \longrightarrow I_{24}$, $I_{31} \longrightarrow -I_{14}$, $I_{34} \longrightarrow -I_{34}$, $I_{14} \longrightarrow -I_{31}$, $I_{24} \longrightarrow I_{23}$ (note that $I_{ij} = -I_{ji}$), and the other 23 substitutions of $G$ are found in the same way from the other 23 permutations of (1, 2, 3, 4). Taking advantage of this partial symmetry, we can write (17) in condensed form as follows:

$$\text{SUMMATION} \left\{ \boldsymbol{440400}_{12} - 4\cdot\boldsymbol{440310}_{24} + 4\cdot\boldsymbol{431400}_{24} + 4\cdot\boldsymbol{430410}_{12} + 6\cdot\boldsymbol{440220}_{12} + 6\cdot\boldsymbol{422400}_{12} + 6\cdot\boldsymbol{420420}_{6} + 12\cdot\boldsymbol{421410}_{24} - 12\cdot\boldsymbol{421401}_{12} - 24\cdot\boldsymbol{411411}_{3} - 4\cdot\boldsymbol{430320}_{24} - 4\cdot\boldsymbol{430230}_{12} - 12\cdot\boldsymbol{431310}_{24} - 4\cdot\boldsymbol{431301}_{24} - 4\cdot\boldsymbol{431130}_{24} + 4\cdot\boldsymbol{431031}_{12} + 4\cdot\boldsymbol{430311}_{24} + 12\cdot\boldsymbol{431220}_{24} + 4\cdot\boldsymbol{430221}_{24} - 12\cdot\boldsymbol{422310}_{24} - 8\cdot\boldsymbol{421320}_{24} + 4\cdot\boldsymbol{421302}_{24} + 12\cdot\boldsymbol{431211}_{24} - 12\cdot\boldsymbol{431121}_{24} + 4\cdot\boldsymbol{421311}_{24} + 6\cdot\boldsymbol{422220}_{24} + 6\cdot\boldsymbol{422022}_{6} + 24\cdot\boldsymbol{422211}_{12} - 12\cdot\boldsymbol{422121}_{24} - 4\cdot\boldsymbol{421221}_{12} + 8\cdot\boldsymbol{421212}_{12} + 40\cdot\boldsymbol{330330}_{3} - 4\cdot\boldsymbol{332310}_{24} + 4\cdot\boldsymbol{332301}_{24} + 44\cdot\boldsymbol{331320}_{24} - 24\cdot\boldsymbol{331311}_{12} + 8\cdot\boldsymbol{332220}_{12} - 4\cdot\boldsymbol{332202}_{24} + 40\cdot\boldsymbol{330222}_{12} + 44\cdot\boldsymbol{322320}_{12} - 4\cdot\boldsymbol{332211}_{24} + 8\cdot\boldsymbol{332112}_{12} + 44\cdot\boldsymbol{331212}_{24} - 64\cdot\boldsymbol{321321}_{6} - 64\cdot\boldsymbol{322221}_{24} - 372\cdot\boldsymbol{222222}_{1} \right\} = 0, \tag{18}$$

with these conventions: instead of, for example, $12a^4b^3cd^2e^2$ we write $12\cdot\boldsymbol{431220}$, setting down only the numerical coefficient and the exponents in order. "*Summation*" is to be understood with respect to the group $G$; that is, the 24 substitutions of $G$ are to be applied to each of the 46 terms appearing in (18), so as to produce the explicit form (17). The 24 substitutions do not always produce 24 distinct terms from a single term of (18); the number actually obtained from the various terms is told by the subscripts in (18). The sum of these subscripts is 805, the total number of terms in the identity. Written in full, the first term of (18) stands for

$$I_{12}^4I_{23}^4I_{34}^4 + I_{12}^4I_{23}^4I_{14}^4 + I_{12}^4I_{31}^4I_{34}^4 + I_{12}^4I_{31}^4I_{24}^4 + I_{12}^4I_{34}^4I_{14}^4 + I_{12}^4I_{34}^4I_{24}^4 + I_{23}^4I_{31}^4I_{14}^4 + I_{23}^4I_{31}^4I_{24}^4 + I_{23}^4I_{34}^4I_{14}^4 + I_{23}^4I_{14}^4I_{24}^4 + I_{31}^4I_{34}^4I_{24}^4 + I_{31}^4I_{14}^4I_{24}^4. \tag{19}$$

What we have proved, then, is this:

THEOREM. *If four curves are tangent to each other at a point (first order contact), they form six horn angles. The six conformal invariants of these horn angles always satisfy the long identity* (17), *which is given explicitly by* (16) *or* (18).

For three curves there is no identity among the $I$'s, but as noted above there is a fundamental inequality, which we state here without proof[6]:

THEOREM. *If three curves are tangent at a point (first order contact), they form three horn angles. The three conformal invariants of these horn angles always satisfy the inequality*

$$I_{12}I_{13} + I_{23}I_{21} + I_{31}I_{32} \geqq 0. \tag{20}$$

Besides the *invariant* $I$ it is convenient to employ its reciprocal, which will be called the *conformal measure* of a horn angle and will be denoted by $M$. Thus by definition,

$$M_{12} = \frac{1}{I_{12}} = \frac{(\gamma_2 - \gamma_1)^2}{d\gamma_2/ds_2 - d\gamma_1/ds_1}. \tag{21}$$

In terms of the conformal measure, the inequality above reduces to a *modified triangle inequality:*

THEOREM. *The conformal measures of the three horn angles formed by three tangent curves satisfy the triangle inequality*

$$M_{12} + M_{23} \geqq M_{13} \text{ if } M_{12}\, M_{23}\, M_{13} > 0, \text{ and} \tag{22}$$

$$M_{12} + M_{23} \leqq M_{13} \text{ if } M_{12}\, M_{23}\, M_{13} < 0. \tag{23}$$

This inequality (with certain sub-conditions[7]) is sufficient as well as necessary for three real numbers to be the conformal measures of the horn angles formed by three tangent curves. This problem and the analogous one for four curves will be discussed later.

If in particular (20) becomes an equality, and hence also (22) and (23), then we shall say the three curves are *in wide open position.* The appropriate condition is obtained, and is shown to be fulfilled for curves obtained by repeated Schwarzian reflections (conformal symmetry). Dual results are valid in equilong geometry.

[1] Abstract in *Bull. Amer. Math. Soc.*, **41**, 784 (1935).

[*] The present work deals with *regular conformal* geometry; and therefore must be distinguished from the classic Riemann problem, in which singular points are allowed on the boundary (see for example the contributions of Caratheodory, Ostrowski and Warschawski; in their work the term cusp is used for a horn angle).

In regular conformal geometry a single regular analytic arc has no invariant (since it can always be converted into a straight line); but an irregular arc with a singularity of the nature of a cusp, in the ordinary sense of algebraic curve theory, does in certain cases possess regular conformal invariants. (See Kasner, *Trans. Amer. Math. Soc.*, **16**, 333(1915).)

[2] E. Kasner, "Conformal Geometry," *Proc. Fifth Internat. Cong. Mathematicians*, Cambridge, **2**, 81 (1912). Invariants of finite order are to be understood here. That is, the series converting one horn angle into another can be formally determined when $I$ is the same, but it may not always converge. Related divergence questions have been studied by G. A. Pfeiffer, *Trans. Amer. Math. Soc.* (1916).

[3] It was shown in the Cambridge reference that there is a unique invariant. For horn angles of second order contact the unique invariant is of fifth order; if the contact is of $k$-th order, the unique invariant is of order $2k + 1$. The total set of horn angles of all contacts form a non-Archimedian system (Veronese, Hilbert, Klein, Levi-Civita). In a book on *Conformal Geometry* being written by E. Kasner, this *non-Archimedian geometry* will be developed systematically with intrinsic conformal methods.

[4] *Jahres. Deut. Math.-Vereinigung*, **44**, 56 (1934).

[5] A. Transon defined the "deviation" of a curve at a point as the tangent of the angle $\delta$ from the axis of the osculating parabola to the normal at the point (*Jour. Mathemat. Pures Appliquées*, **6**, 191 (1841)). The line parallel to the axis of the parabola is Blaschke's affine normal. The fourth expression for $I$ in E. Kasner, "Conformal Geometry," *Sci.*, **82**, 622 (1935) contains a misprint, and should be corrected as above (1′).

[6] See *Bull. Amer. Math. Soc.*, **42**, 181 (1936).

[7] If two of the numbers vanish, so must the third; and the three numbers must not all be equal in absolute value except when they vanish.

---

## ON NORMAL COÖRDINATES

BY T. Y. THOMAS

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY

Communicated April 4, 1936

Let $\mathfrak{M}$ be an $n$-dimensional affinely connected topological space of class one, i.e., the components of the affine connection of $\mathfrak{M}$ are continuous and possess continuous first derivatives.[1] The existence of a system of normal coördinates with origin at an arbitrary point $P$ of $\mathfrak{M}$ was shown by Veblen and Whitehead[1] under the assumption that the relations defining these coördinates possessed continuous first derivatives at $P$. Now the problem of showing the existence of normal coördinates in the space $\mathfrak{M}$ is equivalent to the problem of the existence of such coördinates in a Finsler space under corresponding conditions of continuity and differentiability.[2] For the Finsler space this problem had previously been attacked by W. Mayer[3] who showed that the existence of the above derivatives assumed by Veblen and Whitehead appeared as an essential difficulty on the basis of the usual (geometrical) definition of the normal coördinates. Later Whitehead[4] demonstrated the existence of these derivatives for the case of a Finsler space and so obtained a complete proof of the existence of the normal coördinates. It is one of the objects of the following note to point out that the existence of these coördinates is a direct and immediate consequence of an established theorem on ordinary differential equations.

It is well known that the solution of the system

$$\frac{dx^\alpha}{dt} = u^\alpha, \frac{du^\alpha}{dt} = -\sum_{\mu,\nu=1}^{n} \Gamma^\alpha_{\mu\nu}(x)u^\mu u^\nu \qquad (1)$$

is of the form $x^\alpha = f^\alpha(x_0, \xi t)$ and is uniquely determined by the initial conditions $x^\alpha = x_0^\alpha$ and $dx^\alpha/dt = \xi^\alpha$ for $t = 0$; it follows also from the existence theorem for systems of ordinary differential equations that the functions $f^\alpha(x_0, \xi t)$ satisfy (1) in a domain $|t| \leqq 1$, $|x_0^\alpha - p^\alpha| \leqq B$, $|\xi^\alpha| \leqq B$, where $p^\alpha$ are the coördinates of $P$, in which they are continuous and have continuous first partial derivatives and in fact for values of the variables $t$, $x_0^\alpha$ and $\xi^\alpha$ in this domain the values of the $x^\alpha$ will lie in the domain $|x^\alpha - p^\alpha| \leqq 2B$ where $B$ is a suitable positive constant.[5] Putting $y^\alpha = \xi^\alpha t$ the existence of the derivatives $\partial f^\alpha(x_0, y)/\partial y^\beta$ at $y^\alpha = 0$ follows immediately from the existence of the derivatives $\partial f^\alpha(x_0, \xi)/\partial \xi^\beta$ at $\xi^\alpha = 0$. As ob-

served by Veblen and Whitehead[1] the derivatives $\partial f^\alpha/\partial y^\beta$ at $y^\alpha = 0$ must then have the values $\delta^\alpha_\beta$. Hence the relations $x^\alpha = f^\alpha(x_0, y)$ define normal coördinates $y^\alpha$ for $|x_0^\alpha - p^\alpha| < B$ and $|y^\alpha| < B(x_0) \leqq B$ where $B(x_0)$ is a positive number depending on the coördinates $x_0^\alpha$ of the origin.

For an affinely connected topological space of class $p$ the functions $f^\alpha(x_0, y)$ will be continuous and have continuous derivatives with respect to the variables $x_0^\alpha$, $y^\alpha$ to the order $p$ inclusive and for a space of class $A$ these functions will be analytic in the domain $|x_0^\alpha - p^\alpha| < B$, $|y^\alpha| < B$. The number $B(x_0)$ can be chosen so that in the domain $|y^\alpha| < B(x_0)$ the relations $x^\alpha = f^\alpha(x_0,y)$ defining the normal coördinates $y^\alpha$ will have a unique inverse $y^\alpha = \psi^\alpha(x_0,x)$ such that the functions $f^\alpha(x_0,y)$ and $\psi^\alpha(x_0,x)$ will have the same properties of continuity and differentiability in the variables $y^\alpha$ and $x^\alpha$, respectively. When we speak of the relations $x^\alpha = f^\alpha(x_0, y)$ as defining normal coördinates $y^\alpha$ it is to be understood implicitly that the inverse relations $y^\alpha = \psi^\alpha(x_0, x)$ possess such properties of continuity and differentiability.

Now Whitehead[6] has shown that any point $P$ of $\mathfrak{M}$ is contained in a convex region $R$, i.e., a region $R$ such that any two of its points can be joined by one and only one path which does not leave $R$. Denoting by $p^\alpha$ the coördinates of the point $P$ let us choose the constant $B$ so that the domain $|x^\alpha - p^\alpha| \leqq 2B$ is contained entirely in $R$. Hence $x^\alpha = f^\alpha(x_0, y)$ will have a unique inverse $y^\alpha = \psi^\alpha(x_0, x)$ for $|x_0^\alpha - p^\alpha| < B$, $|y^\alpha| < B$. By suitably restricting $B$ the jacobian $|f^\alpha(x_0, y)/\partial y^\beta|$ will not vanish in the domain $|x_0^\alpha - p^\alpha| < B$, $|y^\alpha| < B$ and hence the functions $\psi^\alpha(x_0, x)$ will be continuous and have continuous first partial derivatives with respect to the variables $x_0^\alpha$ and $x^\alpha$. For values of $x_0^\alpha$ and $y^\alpha$ in the domain $|x_0^\alpha - p^\alpha| < B$, $|y^\alpha| < B$ the relations $x^\alpha = f^\alpha(x_0, y)$ will therefore define a system of normal coördinates.

THEOREM I. *Corresponding to any point $P$ of $\mathfrak{M}$ there exists a constant $B > 0$ and a set of functions $f^\alpha(x_0, y)$ continuous and with continuous first partial derivatives with respect to the variables $x_0^\alpha$ and $y^\alpha$ in the domain $|x_0^\alpha - p^\alpha| < B$, $|y^\alpha| < B$ such that the relations $x^\alpha = f^\alpha(x_0, y)$ define a system of normal coördinates $y^\alpha$ for $|y^\alpha| < B$, the origin being any point with coördinates $x_0^\alpha$ in the domain $|x_0^\alpha - p^\alpha| < B$.*

Strictly speaking the value of the constant $B$ in the above theorem is limited by the coördinates $x^\alpha$ which are selected. In failing to state this dependence and hence in associating the constant $B$ with the point $P$ itself, we adopt the point of view that the conditions of this theorem can be satisfied by making a suitable selection of these coördinates.

Suppose now that $\mathfrak{M}$ is compact. Then the least upper bound of the constants $B$ corresponding to any point $P$ of $\mathfrak{M}$ has a greatest lower bound $b > 0$ in $\mathfrak{M}$. In fact if such were not the case we could find a sequence of points $P_n \longrightarrow P$ such that $B_n \longrightarrow 0$ where $B_n$ denotes the least upper

bound of the constants $B$ for the point $P_n$. But this contradicts Theorem I.

THEOREM II. *There exists a constant $b > 0$ such that for any point $P$ of a compact space $\mathfrak{M}$ the relations $x^\alpha = f^\alpha(x_0, y)$ for $|x_0^\alpha - p^\alpha| < b$, $|y^\alpha| < b$ define a system of normal coördinates.*

We now consider the domain of definition of the normal coördinates $y^\alpha$ in terms of the underlying coördinates $x^\alpha$ of the space $\mathfrak{M}$. In this connection we make use of a trick employed by W. Mayer[7] which consists in adjoining to the relations $x^\alpha = f^\alpha(x_0, y)$ the equations $z^\alpha = x_0^\alpha$. It then follows from the implicit function theorem that the set of $2n$ equations $x^\alpha = f^\alpha(x_0, y)$, $z^\alpha = x_0^\alpha$ has a solution $y^\alpha = \psi^\alpha(z, x)$, $x_0^\alpha = z^\alpha$ uniquely determined by the conditions $y^\alpha = 0$, $x_0^\alpha = p^\alpha$ for $z^\alpha = x^\alpha = p^\alpha$ and such that the functions $\psi^\alpha(z,x)$ are continuous and have continuous first derivatives in the domain $|x^\alpha - p^\alpha| < C$, $|z^\alpha - p^\alpha| < C$. In other words $x^\alpha = f^\alpha(x_0,y)$ has a unique solution $y^\alpha = \psi^\alpha(x_0,x)$ such that $y^\alpha = 0$ for $x_0^\alpha = x^\alpha = p^\alpha$ where the functions $\psi^\alpha(x_0,x)$ are continuous and have continuous first derivatives with respect to the variables $x_0^\alpha$ and $x^\alpha$ in the domain $|x^\alpha - p^\alpha| < C$, $|x_0^\alpha - p^\alpha| < C$. If we take $D = C/2$ a corresponding statement can be made regarding the functions $\psi^\alpha(x_0,x)$ for values of $x_0^\alpha$ in the domain $|x_0^\alpha - p^\alpha| < D$ and for values of $x^\alpha$ in the domain $|x^\alpha - x_0^\alpha| < D$ which is thus determined after the value of $x_0^\alpha$ has been selected. Corresponding to Theorems I and II we now have the following theorem.

THEOREM III. *Corresponding to any point $P$ of $\mathfrak{M}$ there exists a constant $D > 0$ and a set of functions $\psi^\alpha(x_0, x)$ continuous and with continuous first partial derivatives with respect to the variables $x_0^\alpha$, $x^\alpha$ in the domain $|x_0^\alpha - p^\alpha| < D$, $|x^\alpha - x_0^\alpha| < D$ such that the relations $y^\alpha = \psi^\alpha(x_0, x)$ define a system of normal coördinates $y^\alpha$ for $|x^\alpha - x_0^\alpha| < D$, the origin being any point with coördinates $x_0^\alpha$ in the domain $|x_0^\alpha - p^\alpha| < D$. If $\mathfrak{M}$ is compact there exists a constant $d > 0$ such that for any point $P$ of $\mathfrak{M}$ the relations $y^\alpha = \psi^\alpha(x_0, x)$ for $|x_0^\alpha - p^\alpha| < d$, $|x^\alpha - x_0^\alpha| < d$ define a system of normal coördinates $y^\alpha$ with origin at the point with coördinates $x_0^\alpha$.*

In particular $\mathfrak{M}$ may be a compact Riemann space $\mathfrak{M}^*$. Any point P of $\mathfrak{M}^*$ may be made the origin of a system of normal coördinates $y^\alpha$ defined throughout the interior of an $n$-dimensional sphere of radius $\rho$ determined on the basis of the metric of $\mathfrak{M}^*$. In fact

$$\sum_{\alpha,\beta=1}^{n} g_{\alpha\beta}(x_0) y^\alpha y^\beta = \rho^2, \tag{2}$$

for $|x_0^\alpha - p^\alpha| < b$, defines such a spherical surface with center at the point with coördinates $x_0^\alpha$ provided that the value of the radius $\rho$ is such that $|y^\alpha| < b$ where $b > 0$ is the constant specified in Theorem II. Denote by $\sigma$ the least upper bound of the radii $\rho$ for the point $P$. Then $\sigma$ has a greatest lower bound $r > 0$ in $\mathfrak{M}^*$. Otherwise there would exist a sequence of

points $P_m \longrightarrow P$ such that $\sigma_m \longrightarrow 0$ where $\sigma_m$ is the value of the quantity $\sigma$ for the point $P_m$. But this is not possible on account of the continuity of the coefficients $g_{\alpha\beta}(x_0)$ in the above equations (2). The following theorem is therefore established.

THEOREM IV. *If $\mathfrak{M}$ is a compact Riemann space there exists a constant $r > 0$ such that any point $P$ of $\mathfrak{M}$ can be made the origin of a system of normal coördinates defined throughout the interior of an $n$-dimensional sphere of radius $r$ having its center at the point $P$ and determined on the basis of the metric in $\mathfrak{M}$.*

I have been informed by Professor W. Mayer that he has given a direct proof of this latter theorem without appeal to the convexity theorem of Whitehead and that his result will appear in a forthcoming paper on the calculus of variations.

[1] This implies that the coördinate relationships between the coördinates of intersecting neighborhoods are continuous and have continuous partial derivatives to the order three inclusive. Spaces of class $p$ ($> 1$) may be defined in an analogous manner. In an affinely connected topological space of class $A$ the components of the affine connection and hence the above coördinate relationships are analytic. Cf. O. Veblen and J. H. C. Whitehead, "A Set of Axioms for Differential Geometry," *Proc. Nat. Acad. Sci.*, **17**, 551–561 (1931).

[2] For the case of the Finsler space the right members of the second set of equations (1) are replaced by functions $H^\alpha(x,u)$ which are continuous and have continuous derivatives in the variables $x^\alpha$ and $u^\alpha$. The proof of the existence of normal coördinates in the Finsler space is therefore identical with the proof of the existence of such coördinates in the space $\mathfrak{M}$.

[3] A. Duschek-W. Mayer, *Lehrbuch der Differentialgeometrie*, **II**, 95 (1930).

[4] J. H. C. Whitehead, "On the Covering of a Complete Space by the Geodesics through a Point," *Ann. Math.*, **36**, 680 (1935). The assumption of analyticity made by Whitehead appears to be unnecessary in his demonstration so that he has in reality proved the existence of normal coördinates in a Finsler space under the conditions of continuity and differentiability of the functions $H^\alpha(x,u)$ as above stated.

[5] This theorem can be considered as a special case of the existence theorem proved by W. Mayer and T. Y. Thomas, "Vollständig integrable systeme totaler Differentialgleichungen," *Math. Zeit.*, **40**, 658–661 (1936).

[6] J. H. C. Whitehead, "Convex Regions in the Geometry of Paths," Quarterly *Jour. Math.*, **3**, 33–42 (1932).

[7] A. Duschek-W. Mayer, loc. cit., p. 96.

# FUNCTIONAL TOPOLOGY AND ABSTRACT VARIATIONAL THEORY

BY MARSTON MORSE

INSTITUTE FOR ADVANCED STUDY

Communicated April 4, 1936

This paper is an abstract of lectures on "Analysis in the Large" given before a seminar at the Institute for Advanced Study, starting in September, 1935. The author wishes to acknowledge the many helpful comments by the members of the seminar, particularly those by Baer, Čech and Pitcher. See Theorem 1.3.

We are concerned with a function $F$ which is lower semi-continuous on an abstract metric space $M$. Although minima of lower semi-continuous functions have been studied on spaces such as $M$, there has been no adequate investigation of the whole topology of $F$ and $M$. We give three definitions of topological critical points—*homotopic*, *combinatorial* and *essential*, and show that in general these definitions of critical points are distinct from each other and from the ordinary definition of differential critical points when the latter points exist. We give a treatment in which infinite values of the function are on an equal footing with the finite values so that our existence theorems cannot be denied by an escape to infinity. We finish by developing an abstract variational theory in which a fundamental theorem is that a topological extremal is always a metric extremal (that is, the counter part of an ordinary extremal). This enables us to pass directly from the general functional topology to the calculus of variations without the various theorems on approximation and equivalence that have been used so far.

In the author's opinion the most novel and final part of this paper is the introduction of an underlying abstract group theory. To give a brief preliminary indication of this, let it be understood that we are dealing throughout with additive abelian operator groups, in general infinite, with operators $\delta$ in a field, and that all subgroups and isomorphisms are operator subgroups and isomorphisms. The author's earlier infinite set of relations had the form

$$\mu_n - \mu_{n-1} + \ldots + (-1)^n \mu_0 \geqq 0 \qquad [n = 0, 1, \ldots] \qquad (1.1)$$

where $\mu_k = M_k - R_k$, $M_k$ being the $k$th type number and $R_k$ the $k$th connectivity of $M$. The numbers $\mu_k$ were supposed finite.

In the present theory we are concerned with an infinite sequence of groups $m_k$ of $k$-caps, $k = 0, 1, \ldots$ together with certain natural isomorphisms which will be termed *-isomorphisms. The relations (1.1) are replaced by a set of relations of the form

$$m_n \bmod^* m_{n-1} \bmod^* \ldots \bmod^* m_0 \geqq 0 \quad [n = 0, 1, \ldots] \quad (1.2)$$

where $\geqq 0$ is to be read "exists" and mod* $A$ is to be read "mod a group *-isomorphic with $A$." The operations indicated in (1.2) are to be successively performed starting from the right. When the dimensions $\mu_k$ of the groups $m_k$ are finite, the relations (1.1) are a consequence of relations (1.2). Relations (1.2) now center in the *-isomorphisms. When these *-isomorphisms are more explicitly defined, *the fundamental problem in the large is to find conditions under which an infinite sequence of groups $m_k$ of k-caps satisfy* (1.2).

1. *The Underlying Group Theory.*—Let $G$ be an admissible group as defined above. Let $(r)$ be a simply ordered set. With some but not all elements $u$ of $G$ we suppose there is uniquely associated an entity $r$ in $(r)$. We term $r = r(u)$ the *rank* of $u$. The null element in $G$ shall have no rank. In general 0 and the elements of $G$ with rank will not form a group. We suppose these ranks satisfy the following hypotheses.

I. $r[\delta u] = r[u]$ for $\delta \neq 0$. II. If $r(u)$, $r(v)$ *and* $r(u + v)$ *exist,* $r(u + v) \leqq max\,[r(u), r(v)]$. III. *If* $r(u)$ *and* $r(v)$ *exist and are different,* $r(u + v)$ *exists.* IV. *If* $u_1, \ldots, u_m$ *and* $v_1, \ldots, v_n$ *have ranks at most* $r_0$ *and if* $u = \Sigma u_i$ *and* $v = \Sigma v_j$ *have no rank, while* $u + v$ *has a rank, then* $r(u + v) < r_0$.

By a group $g$ of elements of $G$ with property $A$ is meant a group every element of which with the possible exception of 0 has the property $A$. The group $g$ will be termed *maximal* if it is a proper subgroup of no subgroup of $G$ with property $A$. By a *maximal linear set* of elements with property $A$ is meant a maximal set of elements of which every proper linear combination with coefficients in the field has the property $A$. It is readily proved that the cardinal number $h$ of a maximal linear set of non-null elements in an admissible group $g$ is uniquely determined by $g$. We term $h$ the dimension of $g$.

We shall say that two elements $u$ and $v$ with rank are in the same *r-class* if their ranks are the same, while their difference has no rank or has a lesser rank. An isomorphism between two subgroups of $G$ of elements with rank will be termed an *r-isomorphism* if corresponding elements are in the same $r$-class. There always exist maximal groups of elements with rank, but in general these groups will not be $r$-isomorphic, or in fact have the same dimension. Against this we set the following theorems concerning subgroups of $G$.

THEOREM 1.1. *Any two maximal groups of elements with a fixed rank r are r-isomorphic.*

THEOREM 1.2. *Let g be a group of elements with ranks forming a well-ordered set. The group g is a direct sum of arbitrary maximal subgroups of g with the respective ranks of elements of g.*

THEOREM 1.3. *Let g be a group of elements with ranks with a countably*

*infinite dimension. The group g is a direct sum of suitably chosen maximal subgroups of g with the respective ranks of elements of g.*

Theorem 1.3 is due to Baer. The maximal sets in Theorem 1.3 are in general not arbitrary as examples would show.

THEOREM 1.4. *If the set of all possible ranks r is well ordered, any two maximal subgroups of G with rank are r-isomorphic.*

Corresponding to a given type of isomorphism termed a *-isomorphism, we shall denote a group *-isomorphic with a group $g$ by $g^*$. An infinite sequence of groups $m_k$, $k = 0, \ldots$ will be termed **-telescopic* if of the form $m_k = g_k + g^*_{k-1}$, where $g_{-1} = 0$.

THEOREM 1.5. *A necessary and sufficient condition that relations of the form* (1.2) *hold is that the sequence of groups $m_0, m_1, \ldots$ be *-telescopic.*

2. *Natural Isomorphisms.*—We suppose that the function $F$ is lower semi-continuous on $M$, that $0 \leqq F \leqq 1$, and that the sets of points on which $F \leqq c$ are compact when $c < 1$. We do not assume $M$ compact. We obtain generality in this way. In fact if $f$ is a function lower semi-continuous on $M$ with an absolute minimum $m$, but admitting ideal values $+\infty$ we can set $\varphi = f - m$ and $F = \phi/(1 + \phi)$ understanding thereby that $F = 1$ when $f = +\infty$. We thus obtain an admissible function $F$. In this way, for example, we can admit infinite lengths and areas in the applications.

The set of points at which $F < c$ will be denoted by $F_c$. If $K$ is an arbitrary point set we shall say that $K$ is "definitely on" $F_c$ (written on° $F_c$) if on $K$, $F < c' < c$. If $A$ is an arbitrary point set we shall say that $K$ is *on°* $A + F_c$ if $K = K' + K''$ with $K'$ on $A$ and $K''$ on° $F_c$. *Mod°* $F_c$ shall mean mod some compact subset on° $F_c$.

We employ an extension of the Vietoris topology. One starts with the oriented cells of norm $e$ and combines them linearly with coefficients in the field $\delta$, thus forming "algebraic chains of norm $e$." Our $k$-cycles are then defined as are the "fundamental sequences" of Vietoris. Let $e_n$ be a sequence of positive numbers which tend to zero as $n$ becomes infinite. Let $A$ and $C$ be compact sets such that $A$ is on $C$. Let $v$ be a sequence of algebraic $k$-chains $v_n$ of norm $e_n$, such that for $n = 0, 1, \ldots, v_n \overset{e_n}{\sim} v_{n+1}$ mod $A$ on $C$. Then $v$ is termed a $k$-cycle mod $A$ on $C$ and $C$ a *carrier* of $v$. If for a sequence $e'_n$ similar to $e_n$, $v_n \overset{e'_n}{\sim} 0$ mod $A$ on $C$, we say that $v \sim 0$ mod $A$ on $C$. Carriers of cycles or homologies will always be understood to be compact. We shall say that a relative $k$-cycle or homology is on a domain $H$ if the cycle or homology admits a carrier on $H$.

If $z$ is a $k$-cycle on° $F_b$ not homologous to zero on° $F_b$, we term $b$ a homology limit of $z$ and $z$, $F$-non-bounding. By the *superior homology limit* $s(z)$ of an $F$-non-bounding $k$-cycle is meant the least upper bound of the homology limits of $z$. If $z$ is non-bounding on $M$ we understand that $s(z) = +\infty$. Let $z$ be a $k$-cycle with superior homology limit $s(z) = b$.

By the *inferior homology limit* $t(z)$ of $z$ is meant the greatest lower bound of constants $c$ such that $z \sim 0$ on° $F_b$ mod° $F_c$.

If $z$ is an $F$-non-bounding $k$-cycle we term the pair $[s(z), t(z)]$ the *rank* $r(z)$ of $z$ and order these pairs lexigraphically.

THEOREM 2.1. *The ranks $r(u)$ of $F$-non-bounding $k$-cycles $u$ satisfy* I, II, III, IV, *the group $G$ being the group of all $k$-cycles.*

This theorem is fundamental.

A sequence of algebraic $k$-chains $w_n$ with norms $e_n$ tending to zero will be called a *formal $k$-chain $w$*. To add two formal $k$-chains we add their components. A $k$-cycle $u$ mod° $F_a$ on $F \leqq a$ will be termed a *$k$-cap* belonging to $a$ if there exists a constant $b > a$ such that $u \frown\!|\!\sim 0$ on° $F_b$ mod° $F_a$. If $u$ is a $k$-cap belonging to $a$ we term $a$ a *cap limit* $a(u)$. To add two $k$-caps $u$ and $v$ we first add $u$ and $v$ as formal $k$-chains. If the components of $u + v$ define a $k$-cap $w$ belonging to some constant $c$, we show that $c$ is uniquely determined by $u$ and $v$ and say that $u$ and $v$ sum to the $k$-cap $w$. If no such constant as $c$ exists, we say that the formal $k$-chain $u + v$ is not a $k$-cap. The cap limits $a(u)$ satisfy I, II, III, IV, the underlying group being the group of formal $k$-chains.

Two $k$-caps belonging to $a$ will be said to belong to the *same $a$-class* if their difference is not a $k$-cap belonging to $a$. An isomorphism between two groups of $k$-caps in which corresponding $k$-caps are in the same $a$-class, will be termed a *cap-isomorphism*.

A $k$-cap $u$ such that $Bu \frown\!|\!\sim 0$ on° $F_{a(u)}$ will be said to *cap* any $(k-1)$-cycle in the $r$-class of $Bu$. A $k$-cycle which is also a $k$-cap will be called a *cycle-cap*. We show that there is a cycle-cap in the $r$-class with each $F$-non-bounding $k$-cycle. A group $g_k^*$ of $(k+1)$-caps will be said to be **-isomorphic* with a group $g_k$ of $k$-cycles if $g_k^*$ and $g_k$ admit an isomorphism in which each cycle in $g_k$ is capped by its correspondent in $g_k^*$. A group $g_k$ of $F$-non-bounding $k$-cycles which is a direct sum of subgroups of $g_k$ of the different ranks of $g_k$ will be termed *neuclear*.

THEOREM 2.3. *Any neuclear group $g_k$ of $k$-cycles of finite rank is *-isomorphic with at least one group $g_k^*$ of $(k+1)$-caps. If $g_k$, $k = 0, 1, \ldots$ is an arbitrary sequence of neuclear groups of cycle-caps with finite ranks, the direct sums $m_k = g_k + g_{k-1}^*$ $k = 0, 1, \ldots$, are permissible and define a *-telescopic sequence of groups of $k$-caps which satisfy* (1.2).

The fundamental problem in the large enunciated in §1 admits the following two solutions.

*Definition.*—A $k$-cap is termed conditional if there are no $k$-cycles in its $a$-class or at least one conditional $k$-cycle.

THEOREM 2.4. *Let there be given an arbitrary sequence of groups $M_k$ of conditional $k$-caps.*

(1) *If dim. $M_k$ is finite for each $k$, there exists a *-telescopic sequence of*

*groups $m_k$ of finite dimension such that $M_k$ is cap-isomorphic with a subgroup of $m_k$.*

(2) *If the ranks of all F-non-bounding k-cycles are well-ordered and the groups $M_k$ are maximal, the groups $M_k$ are cap-isomorphic with groups of a *-telescopic sequence in which the groups $g_k$ are maximal.*

*In applying* (2) *we can make use of the fact that the groups $M_k$ in* (2) *are of the form $M_k = N_k \bmod R_k$* where $N_k$ is a maximal group of $k$-caps and $R_k$ is a maximal group of unconditional caps. $R_k$ is isomorphic with the $k$th homology group of $M$ and in this way the connectivities enter into our relations. In particular the numbers $u_k$ = dim. $N_k$ − dim. $R_k$, if finite, satisfy the relations (1.1).

3. *Homotopic Critical Points.*—Let $E$ be a compact subset of $M$. We consider deformations $D$ of $E$ which replace a point $p$ on $E$ when $t = 0$ by a point $q = f(p, t)$ on $M$ at the time $t$. We suppose that $f(p, t)$ is continuous in $p$ and $t$ for $p$ on $E$ and $t$ on an interval of the form $0 \leqq t \leqq t_0$. We say that $D$ admits a displacement function $h(e)$ if for any two points $q_1$ and $q_2$ appearing on the trajectory of a point of $E$ in the order $q_1, q_2$ with the distance $q_1q_2 > e$, $F(q_1) - F(q_2) > h(e)$ where $h(e)$ is a positive function of $e$. If $D$ admits a displacement function we term $D$ an *F-deformation*. We term $D$ *proper* if on none of its trajectories is the point $q$ fixed for a time interval of the form $[0, t_1]$.

If $p$ is a point on $M$ and $c$ a value between 0 and 1 inclusive, we term a pair $(p, c)$ admissible if $c = F(p)$ or if $c$ is a limiting value of $F$ at $p$. We set $\zeta = (p, c)$ and term the set of admissible pairs $(p, c)$ the space $(\zeta)$. If $H$ is a set of pairs $\zeta$ in the space $(\zeta)$ we shall term the corresponding set of points $p$ the $p$-projection of $H$ and denote it by $H^p$. A pair $\zeta$ will be termed *homotopically ordinary* if corresponding to an arbitrary compact set $C$ on $M$ there exists a neighborhood $V$ of such that the set $C \cdot \overline{V^p}$ on $M$ admits a proper $F$-deformation. A pair which is not homotopically ordinary will be termed *homotopically critical.*

When $F$ is a differentiable function of coördinates, a homotopic critical pair $(p, c)$ yields a differential critical point $p$ but not conversely, as the example $x^3$ shows. An important property of critical pairs is that they are closed on the space $(\zeta)$.

THEOREM 3.1. *If $c$ is a finite superior or inferior homology limit or a cap limit there is at least one critical pair of the form $(p, c)$.*

If $(p, c)$ is a critical pair we term $p$ a critical point corresponding to $c$ (written corr. $c$). By a *complete critical set corr. $c$* we mean the set of all points $p$ in critical pairs $(p, c)$. By a *critical* set $\sigma$ corr. $c$ we mean a sum of components (point set sense) of the complete critical set corr. $c$. We say that a $k$-cap $u$ belonging to $c$ *belongs* to $\sigma$ if there is a $k$-cap in the $a$-class of $u$ in every neighborhood of $\sigma$. We show that every $k$-cap belonging to $c$ belongs to some critical set. A critical set $\sigma$ which possesses at least one

$k$-cap will be termed *$k$-essential* if there exists no proper critical subset to which all of its $k$-caps belong.

THEOREM 3.2. *Every critical set $\sigma$ with $c < 1$ to which at least one $k$-cap belongs, possesses a $k$-essential critical subset.*

A critical pair $(p, c)$ such that $p$ belongs to an essential critical set $\sigma$ corr. $c$ will be termed *essential*. We show by example that there exist homotopic critical pairs which are not essential.

4. *Combinatorial Critical Points.*—We term $b$ an upper bound of a relative cycle $u$ if $u$ lies on $F \leqq b$. A $k$-cycle $u$ mod $A$ on $C$ will be said to be *sub-homologous* to zero on $C$ if the following is true. Corresponding to an arbitrary upper bound $b$ of $u$ and arbitrary carrier $K$ of $Bu$ there exists a $k$-cycle $v$ mod $A$ on $C$ such that $u \sim v$ mod $A$ on $C$ and on $F \leqq b$, while $v$ admits a carrier on which $F$ is bounded away from $b$ on any subset which excludes a neighborhood of $K$. A set $S$ of relative $k$-cycles will be said to be *locally sub-homologous* to zero if corresponding to an arbitrary positive constant $e$ there exists a positive constant $\delta$ such that every relative $k$-cycle of $S$ of diameter at most $\delta$ is sub-homologous to zero on a set $C$ of diameter less than $e$.

An admissible pair $\zeta$ will be termed *combinatorially $k$-ordinary* if corresponding to any compact set $C$ on $M$ there exists a neighborhood $V$ of $\zeta$ in the space $(\zeta)$ such that relative $k$-cycles on $C \cdot \overline{V^p}$ are locally sub-homologous to zero. Pairs $\zeta$ not combinatorially $k$-ordinary are termed *combinatorially $k$-critical*.

Combinatorial critical pairs $(p, c)$ are not closed, but for a fixed $c$ they are closed. We show that homotopic critical pairs are combinatorial critical pairs. The converse is not true as we show by example.

THEOREM 4.1. *If $c$ is a finite superior or inferior homology limit of a $k$-cycle or a $k$-cap limit, there is at least one combinatorial critical pair of the form $(p, c)$.*

Our definitions and affirmations with regard to homotopic critical sets $\sigma$ hold for the combinatorial case as well.

5. *Abstract Variational Theory.*—We start with a metric space $N$ with metric in general unsymmetric. Let $|pq|$ denote the maximum of $pq$ and $qp$. The numbers $pq$ define an ordinary symmetric metric on $N$. To avoid ambiguity we refer to $N$ with the metric $|pq|$ as $|N|$. By a parameterized curve $\lambda$ (written par-curve) we mean a continuous image on $|N|$ of a line segment $0 \leqq t \leqq a$. The curve $\lambda$ lies on $N$ and *we use the metric of $N$ to define the length $f$*, possibly infinite, of $\lambda$.

If $h$ and $k$ are two par-curves, the Fréchet distance between $h$ and $k$ on $|N|$ will be denoted by $hk$. The curves $h$ and $k$ will be said to belong to the *same class* if $hk = 0$. We shall speak of a class of par-curves as a *curve*.

Let $a$ and $b$ be two points of $|N|$ and $M(a, b)$ the space of curves joining $a$ to $b$ on $|N|$ assigning $M(a, b)$ the Fréchet metric. If $f$ is the length of

curves of $M(a, b)$, we define $F$ on $M$ as in §2. From this point on, we assume that $|N|$ is compact and connected.

THEOREM 5.1. *For $0 \leqq c < 1$ the subspace $F \leqq c$ of $M(a, b)$ is compact. If $|N|$ is arc-wise connected, the kth homology group of any two spaces $M(a, b)$ and $M(a', b')$ are isomorphic.*

The theory of the preceding sections can now be applied to $M(a, b)$ and the function $F$.

A simple arc on $|N|$ joining $p$ to $q$ will be called *elementary* provided a point $r$ is on the arc if and only if $pq = pr + rq$. We now make the following hypothesis.

*Hypothesis of Local Convexity.—There exists a positive constant $\rho$ such that when $0 < pq \leqq \rho$ there is an elementary arc $pq$ joining $p$ to $q$.*

We show that $|N|$ is arc-wise connected and that $M(a, b)$ is uniformly locally arc-wise connected, and separable. On $M(a, b)$ each inferior homology limit is less than 1, and the dimension of a maximal group of $k$-cycles with a given rank is finite.

THEOREM 5.2. *Each critical pair $(p, c)$ on $M$ is of the form $[p, F(p)]$.*

We accordingly now term a point $p$ in a critical pair a critical point and the corresponding curve on $N$ an *extremal, homotopic, combinatorial or essential*, according to the classification of $p$. We term an extremal of any one of these three types topological and prove the following basic theorem.

THEOREM 5.2. *A topological extremal is a metric extremal.*

In particular if $N$ is any locally extremal convex metric defined for the points of an $n$-sphere $n > 1$, we show that there are infinitely many metric extremals joining any two fixed points.

[1] Birkhoff, G. D., and Hestenes, M., "Generalized Minimax Principle in the Calculus of Variations," *Duke Math. Jour.*, **1**, 413–432 (1935).

[2] Hestenes, M., "Minimax Principle for Functions," *Proc. Nat. Acad. Sci.*, **22**, 115–119 (1936).

[3] Lefschetz, S., "On Critical Sets," *Duke Math. Jour.*, **1**, 392–412 (1932). See also his Coll. Lect. on Topology for his theory of relative cycles.

[4] Menger, K., "Untersuchungen über allgemeine Metrik," *Math. Ann.*, **100**, 75–163 (1928); **103**, 466–501 (1930); "Metrische Geometrie und Variationsrechnung," *Fund. Math.*, **25**, 441–458 (1935). There are deep theorems on the existence of minima to be found here.

[5] Morse, M., "Calculus of Variations in the Large," *Am. Math. Soc., Coll. Publ.*, **18**.

[6] Morse, M., and van Schaack, G., "Abstract Critical Sets," *Proc. Nat. Acad. Sci.*, **21**, 258–262 (1935). Homotopic critical points are first introduced here.

[7] Vietoris, L., "Ueber den höheren Zusammenhang kompakter Räume," *Math. Ann.*, **97**, 454–472 (1926–27).

[8] References to Čech will be given in the complete paper, for certain useful theorems on Vietoris cycles which we have not quoted here.

# A THEORY OF CLASSES PRESUPPOSING NO CANONS OF TYPE

BY W. V. QUINE

SOCIETY OF FELLOWS, HARVARD UNIVERSITY

Communicated April 8, 1936

Starting with a system whose elements are classes or attributes of lowest type, it is proposed to develop the higher types by definition. Expressions purporting to denote higher-type classes emerge as properly non-denotative segments of formulae about classes of lowest type; they are incomplete symbols, defined in context. Only such locutions issue from these definitions as would be countenanced by the familiar theory of types. The theory of types itself, as an *index expurgatorius*, therefore becomes superfluous;[1] the forms which it brands meaningless are already meaningless here in the strict sense of not being defined in terms of the primitives. The system is so fashioned incidentally as not to involve use of variables or parentheses.

Strictly, the intention is to repudiate classes entirely and to regard a term as denoting, not a class of objects, but the several objects themselves. The more familiar idiom used above will be retained for brevity in this synopsis, but translation of the whole into the stricter idiom presents no essential difficulty.

Only the primitives and definitions will be presented here. Argument will be rested intuitively upon the meanings assigned to the signs. Consistent postulates can presumably be framed to fit these meanings, since the paradoxes which arise from confusion of logical type are inexpressible in these terms. Apart from those paradoxes, indeed, certain semantic paradoxes such as Grelling's are to be apprehended in any system which involves semantic ingredients as this one does;[2] but these paradoxes seem likewise to resist translation into terms of the present primitives.

Imagine first a language, comprehensive like English, satisfying these conditions: (I) Each word ends in "b" and consists otherwise solely of "a"s. (The several letters in a word are integral parts of a simple sign.) (II) Each word is of just one of the following kinds: a term, i.e., a class name; a unitary operator, i.e., a word whose prefixture to a term produces a term; a binary operator, i.e., a word whose prefixture to two juxtaposed terms produces a term; and so on. (III) Each operator is a word, and each term which is not a word consists of an $n$-ary operator, for some $n$, followed by $n$ terms.

The grouping of letters into words needs no parenthesis or spacing for its indication in any context, for the ends of words are marked by the "b"s. Since each operator precedes its arguments and is fixed as to the number of admissible arguments, the significant grouping of words is likewise deter-

minate without help of parentheses.[3] All contexts appear thus as uninterrupted strings of "a"s and "b"s.

By an *expression thing* I shall mean an expression as a single physical object.[4] By an *expression class* I shall mean the class of all the expression things which have a given shape—which are "replicas"[5] of one expression. More specifically I shall speak also of term things, term classes, letter things and letter classes. Single quotes will refer to expression classes; thus 'a' is the class of all things shaped like the quoted sample. Double quotes, and the unqualified words "expression," "term," etc., are reserved for the more slipshod idiom, which will for brevity still be used wherever possible.

Five words of the hypothetical language, "b," "ab," "aab," "aaab" and "aaaab," will now be given meaning. The first is to be a term, and is to denote the class of all terms of the hypothetical language itself. (It results in particular that each letter thing "b" is a member of b, and that 'b' is a subclass of b.) The other four words are to be operators, two unitary and two binary, as follows. Where $X$ is any class, ab$X$ is the class of all term things which denote $X$. Where $X$ and $Y$ are any classes, aaaab$XY$ is the class 'a' or the class 'b' according as $X$ is or is not included in $Y$. As a means of describing the other two operators, a thing $x$ will be called the *concatenate* of $y$ upon $z$ if $x$ is a row of letter things, spaced and aligned in standard typographical fashion, $y$ is a proper initial segment of $x$, and $z$ is the rest of $x$. Now aaab$XY$ is the class of all concatenates of members of $X$ upon members of $Y$; and aab$X$ is the class of all things $z$ such that, for some member $x$ of $X$, $z$ is the concatenate of $x$ upon a term thing shaped like (i.e., of the same term class as) a term thing which is an initial segment (or all) of $x$. The variables used expositorily here have no place in the language; operators will there occur in application only to constant terms.

Such are the five primitives of the logico-semantic system occupying this paper. In abstracting this system from the rest of the hypothetical language, let us remember that the class b still embraces as members the term things not only of this sublanguage but of the whole hypothetical language. Similar scope will be intended in any general reference to terms, operators, etc.

Definitions are viewed here as conventions extraneous to the formal system; each formula is to be thought of officially as expanded into full primitive expression. A definition is thus legitimate if and only if unambiguously eliminable.

Since the class b of term things is included in itself but not in the class 'a', the explanation of the operator "aaaab" shows the letter classes to be definable thus:

'a' $=_{df}$ aaaabbb. 1.1. 'b' $=_{df}$ aaaabb'a'. 1.2.

Then aaab'a''b', being the class of concatenates of members of 'a' upon members of 'b', will be the expression class 'ab'; aaab'ab''a', in turn, will be 'aba'; and in this fashion the single quotes are definable for all rows of "a"s and "b"s, as epitomized by these recursive clauses:

$$\text{'}\mathbf{X}\text{a'} =_{df} \text{aaab'}\mathbf{X}\text{''a'}. \quad 1.3. \qquad \text{'}\mathbf{X}\text{b'} =_{df} \text{aaab'}\mathbf{X}\text{''b'}. \quad 1.4.$$

A bold-face capital, used expositorily as here, is supposed supplanted by various specific expressions, terms or otherwise. Single or double quotes attached to a bold-face capital do not involve reference to that capital as such, but apply to the expression by which that capital is supposed supplanted.

Though sentences communicate, basically they do not denote;[6] they figure rather as true or false. But I shall assimilate truth and falsehood to denotation by using true sentences and false sentences arbitrarily as names respectively of the class 'a' and the class 'b'. Sentences then do incidental service as terms, and the classes Ver and Fal of true and false sentence things become definable respectively as the class of term things denoting 'a' and the class of term things denoting 'b':

$$\text{Ver} =_{df} \text{ab'a'}. \quad 2.1. \qquad \text{Fal} =_{df} \text{ab'b'}. \quad 2.2.$$

Under this assimilation, terms produced by the operator "aaaab" are sentences; true or false, moreover, according as the class denoted by the first argument is or is not included in that denoted by the second. The operator hence serves as the copula of inclusion, and is translatable into familiar notation thus:

$$(X \subset Y) =_{df} \text{aaaab}XY. \quad 3.1.$$

If $X$ is one of the classes 'a' and 'b', $X$ is included in 'b' if and only if $X$ is 'b'; and if $X$ and $Y$ are 'a' or 'b', the class aaab$XY$ of concatenates of members of $X$ upon members of $Y$ is included in 'aa' if and only if $X$ and $Y$ are both 'a'. In view of the described assimilation, then, the denial and conjunction of sentences can be introduced thus:

$$\sim X =_{df} (X \subset \text{'b'}). \quad 3.4. \qquad (X.Y) =_{df} (\text{aaab}XY \subset \text{'aa'}). \quad 3.5.$$

The other truth functions are then definable in familiar fashion; e.g.,

$$(X \supset Y) =_{df} \sim(X.\sim Y). \quad 3.6.$$

Identity is now definable indifferently as mutual inclusion or as follows:

$$(X = Y) =_{df} (\text{ab}X \subset \text{ab}Y). \quad 3.9.$$

3.1–3.9 introduce operations upon classes. It is not essential to 3.4–3.6 that these classes be 'a' or 'b'; what 3.4–3.6 yield in non-sentential application is merely indifferent to us.

The next developments lead to quantification. Let "**Z**" be a sentence form in "$\alpha$"; i.e., think of this capital as supplanted in what follows by an expression so constituted of "a"s, "b"s and "$\alpha$"s that if "$\alpha$" were replaced by any term the whole would be a sentence, a term of the hypothetical language denoting 'a' or 'b'. A sentence form "**Y**" will be called *formally equivalent* to "**Z**" if every term whose substitution for "$\alpha$" turns "**Z**" into a true sentence turns "**Y**" into one and *vice versa*. Any sentence thing whose shape is got by substituting a term for "$\alpha$" throughout "**Z**" will be called an *alpha-instance of* "**Z**."

"**Z**," if as described, will be a juxtaposition "$\mathbf{Z}_1\mathbf{Z}_2 \ldots \mathbf{Z}_n$" where each "$\mathbf{Z}_i$" is either "$\alpha$" or a term or an operator; and we may take each "$\mathbf{Z}_i$" maximally large, so that no segment "$\mathbf{Z}_i\mathbf{Z}_{i+1} \ldots \mathbf{Z}_j$" ($i < j$) is a term. Now, for each $h$ and $k$ such that $1 \leqslant h \leqslant k \leqslant n$, $[\mathbf{Z}_h\mathbf{Z}_{h+1} \ldots \mathbf{Z}_k]_\alpha$ is to be the class of all expression things having any shape "$\mathbf{Z'}_h\mathbf{Z'}_{h+1} \ldots \mathbf{Z'}_k$" of the following sort: Where "$\mathbf{Z}_i$" is an operator, "$\mathbf{Z'}_i$" is the same operator; where "$\mathbf{Z}_i$" is a term, "$\mathbf{Z'}_i$" is a term denoting the same class, i.e., a term such that $(\mathbf{Z}_i = \mathbf{Z'}_i)$; and there is a term "**T**" such that, wherever "$\mathbf{Z}_i$" is "$\alpha$," "$\mathbf{Z'}_i$" is "**T**." It is readily seen in particular that (1) every alpha-instance of "**Z**" is a member of $[\mathbf{Z}]_\alpha$, i.e., of $[\mathbf{Z}_1 \ldots \mathbf{Z}_n]_\alpha$, and (2) every member of $[\mathbf{Z}]_\alpha$ is an alpha-instance of a sentence form formally equivalent to "**Z**."

Where "$\mathbf{Z}_i$" is not "$\alpha$," the foregoing specifications identify $[\mathbf{Z}_i]_\alpha$ with the class ab$\mathbf{Z}_i$ or the class '$\mathbf{Z}_i$' according as "$\mathbf{Z}_i$" is a term or an operator. (In view of (I) and (III), 1.1–1.4 provide '$\mathbf{Z}_i$' for any operator "$\mathbf{Z}_i$.") Again, if "$\mathbf{Z}_{j+1}$" is not "$\alpha$," $[\mathbf{Z}_i \ldots \mathbf{Z}_{j+1}]_\alpha$ will be aaab$[\mathbf{Z}_i \ldots \mathbf{Z}_j]_\alpha$ $[\mathbf{Z}_{j+1}]_\alpha$. In this fashion, if "$\mathbf{Z}_m$" is the first occurrence of "$\alpha$" in "**Z**," we arrive recursively at a definition of $[\mathbf{Z}_1 \ldots \mathbf{Z}_{m-1}]_\alpha$ involving only our primitives plus whatever foreign terms there may have been among "$\mathbf{Z}_1$," "$\mathbf{Z}_2$," ... and "$\mathbf{Z}_{m-1}$" themselves.

$[\mathbf{Z}_m]_\alpha$ itself will be the class b, since the expression "**T**" supplanting "$\alpha$" was to be any term. Now if we have defined $[\mathbf{Z}_m \ldots \mathbf{Z}_k]_\alpha$, and "$\mathbf{Z}_{k+1}$" is not "$\alpha$," $[\mathbf{Z}_m \ldots \mathbf{Z}_{k+1}]_\alpha$ is forthcoming as aaab$[\mathbf{Z}_m \ldots \mathbf{Z}_k]_\alpha[\mathbf{Z}_{k+1}]_\alpha$ in the previous fashion. If on the other hand "$\mathbf{Z}_{k+1}$" is "$\alpha$," any expression thing belonging to the class $[\mathbf{Z}_m \ldots \mathbf{Z}_{k+1}]_\alpha$ will consist of a member $x$ of $[\mathbf{Z}_m \ldots \mathbf{Z}_k]_\alpha$ followed by a recurrence of that term, initial to $x$, which supplanted "$\mathbf{Z}_m$." That term is *the* initial term of $x$, for no two terms can be initial one to another.[7] Hence the class $[\mathbf{Z}_m \ldots \mathbf{Z}_{k+1}]_\alpha$ is aab$][\mathbf{Z}_m \ldots \mathbf{Z}_k]_\alpha$. In either case, therefore, $[\mathbf{Z}_m \ldots \mathbf{Z}_{k+1}]_\alpha$ is definable given $[\mathbf{Z}_m \ldots \mathbf{Z}_k]_\alpha$. By induction, then, since $[\mathbf{Z}_m]_\alpha$ is definable as b, $[\mathbf{Z}_m \ldots \mathbf{Z}_n]_\alpha$ is definable. $[\mathbf{Z}]_\alpha$ is now definable as aaab$[\mathbf{Z}_1 \ldots \mathbf{Z}_{m-1}]_\alpha[\mathbf{Z}_m \ldots \mathbf{Z}_n]_\alpha$; definable on the basis ultimately of our primitives, therefore, plus whatever foreign terms there may be in "**Z**" itself.

The above method of defining $[\mathbf{Z}]_\alpha$ is embodied, with inessential refinements, in this recursive definitional scheme:

$$[\alpha]_\alpha =_{df} \mathrm{b}. \qquad 4.1.$$

If "**X**" is a term, $[\mathbf{X}]_\alpha =_{df} \mathrm{ab}\mathbf{X}.$ 4.2.

If "**X**" is an operator, $[\mathbf{X}]_\alpha =_{df} \text{`}\mathbf{X}\text{'}.$ 4.3.

If "**X**" begins in "$\alpha$," $[\mathbf{X}\alpha]_\alpha =_{df} \mathrm{aab}[\mathbf{X}]_\alpha.$ 4.4.

If (I) "**X**" contains no "$\alpha$" and "**Y**" beings in "$\alpha$," or if (II) "**Y**" contains no "$\alpha$" and (i) "**X**" begins and ends in "$\alpha$" or (ii) "**X**" is a term or (iii) "**X**" is an operator and "**XY**" does not begin in a term, then:

$$[\mathbf{XY}]_\alpha =_{df} \mathrm{aaab}[\mathbf{X}]_\alpha[\mathbf{Y}]_\alpha. \qquad 4.5.$$

All such definitions in "$\alpha$" are intended to hold also for any other Greek letter.

Now the quantification "$(\alpha)\mathbf{Z}$" tells us that the sentence form "**Z**" is true for all values of "$\alpha$"; i.e., that substitution of any term for "$\alpha$" throughout "**Z**" (hence also throughout any equivalent form) yields a true sentence. In view of (1) and (2) above, this sense is afforded by the following definition. The accompanying definition of particular quantification is familiar:

$$(\alpha)\mathbf{Z} =_{df} ([\mathbf{Z}]_\alpha \subset \mathrm{Ver}). \quad 5.1 \qquad (\exists\alpha)\mathbf{Z} =_{df} \sim(\alpha)\sim\mathbf{Z}. \quad 5.2.$$

Iterated quantification reduces through 5.1 to iteration of the form "$[\ldots[\ldots]_\alpha\ldots]_\beta$," and thus requires explanation of "$[\mathbf{Z}]_\alpha$" where "**Z**" contains many Greek letters. This is accomplished by handling all Greek letters but "$\alpha$" as terms while applying 4.1–4.5 to "$[\mathbf{Z}]_\alpha$"; "$[\mathbf{Z}]_\alpha$" then reduces to an expression containing the other Greek letters but not "$\alpha$," and we are ready to apply 4.1–4.5 to the next broader context. (This treatment would have been impossible if 4.2 were framed like 4.3.)

Where "**Z**" is any sentence form of the hypothetical language, any result of quantifying "**Z**" with respect to all Greek letters therein contained is reducible through the above definitions to our primitives plus whatever further terms there are in "**Z**." If "**Z**" is built up wholly of our primitives, plus the Greek letters, the quantification reduces wholly to our primitives.

The function $(X \wedge Y)$ or $-X \frown -Y$ of the ordinary class algebra is now definable contextually thus:

The result of putting "$(\mathbf{X} \wedge \mathbf{Y})$" for "$\alpha$" throughout "**Z**" will stand for

$$\text{"}(\alpha)((\beta)((\gamma)(\delta)((((\beta \subset \gamma).(\mathbf{X} \subset \gamma)).(\mathbf{Y} \subset \gamma)) \supset (\delta \subset \gamma)) = (\alpha \subset \beta)) \supset \mathbf{Z}).\text{"} \qquad 6.1.$$

(Certain stipulations, wanted for unique eliminability, are suppressed here.) The rest of the class algebra is thence constructible with help of 3.1–3.9.[8]

Let Greek letters with subscripts serve as supplementary Greek letters, and let this abbreviation be adopted in the use of operators:

$$(\text{aaab})^1 = {}_{\text{df}}\ \text{aaab}. \quad 7.1. \qquad (\text{aaab})^{n+1} = {}_{\text{df}}\ \text{aaab}(\text{aaab})^n. \quad 7.2.$$

The introduction of classes of the higher types will depend in part upon this *principle of regimentation:* For every sentence form "**Z**" whose Greek letters are "$\alpha_1$," "$\alpha_2$," ... and "$\alpha_n$" there are classes $X_1$, $X_2$ ... and $X_{n+1}$ such that "**Z**" is formally equivalent to the sentence form which the definitions 1.1–7.2 abbreviate as

"$((\text{aaab})^{2n} X_1 \text{ab} \alpha_1 X_2 \text{ab} \alpha_2 \ldots X_{n+1} \subset \text{Ver})$."

To establish this we observe that "**Z**" is equivalent to

"$(\beta_1)(\beta_2)\ldots(\beta_n)((\ldots((\beta_1 = \alpha_1).(\beta_2 = \alpha_2))\ldots.(\beta_n = \alpha_n)) \supset \mathbf{Z}')$"

where "**Z**'" is "**Z**" with each "$\alpha$" rewritten as "$\beta$." Given "**Z**," 1.1–5.1 explain the above formula as an abbreviation ultimately of a certain sentence form "$\mathbf{Y}_1\alpha_1\mathbf{Y}_2\alpha_2\ldots\mathbf{Y}_{n+1}$" whose "$\mathbf{Y}_i$" contain only "a"s and "b"s. Since "$\mathbf{Y}_1\alpha_1\mathbf{Y}_2\alpha_2\ldots\mathbf{Y}_{n+1}$ is equivalent in turn to

"$((\text{aaab})^{2n}$'$\mathbf{Y}_1$' ab$\alpha_1$ '$\mathbf{Y}_2$'ab$\alpha_2\ldots$'$\mathbf{Y}_{n+1}$' $\subset$ Ver),"

the classes '$\mathbf{Y}_i$' as of 1.1–1.4 qualify as the classes $X_i$ whose existence was to be proved.

The form of notation "$\alpha \ni \mathbf{Z}$," purporting to denote the class of all classes satisfying the condition "**Z**," is introduced contextually thus:

$$(X \in \alpha \ni \mathbf{Z}) =_{\text{df}} (\alpha)((X = \alpha) \supset \mathbf{Z}). \qquad 8.1.$$

The following definitions extend the devices of 3.1, 3.9 and 6.1 to these ostensible symbols of classes of classes:

$$(\mathbf{X} \subset \mathbf{Y}) =_{\text{df}} (\alpha)((\alpha \in \mathbf{X}) \supset (\alpha \in \mathbf{Y})). \qquad 8.4.$$

$$(\mathbf{X} = \mathbf{Y}) =_{\text{df}} (\alpha)((\alpha \in \mathbf{X}) = (\alpha \in \mathbf{Y})). \qquad 8.5.$$

$$(\mathbf{X} \wedge \mathbf{Y}) =_{\text{df}} \alpha \ni (\sim(\alpha \in \mathbf{X}).\sim(\alpha \in \mathbf{Y})). \qquad 8.6.$$

The whole class algebra is thereby reproduced in the second type.

For such classes of classes, variables "$\alpha^2$," "$\beta^2$," etc., will now be introduced. Definition of quantification with respect to these variables is to be such as to make "$(\alpha^2)\mathbf{Z}$" true if and only if, for every possible sentence form "**W**" of our unlimited hypothetical language, containing "$\beta$" as sole Greek letter, the result of putting "$\beta \ni \mathbf{W}$" for "$\alpha^2$" in "**Z**" is true. For each "**W**" there are, by the principle of regimentation, classes $X_1$ and $X_2$ such that

$$(\beta \ni \mathbf{W} = \beta \ni ((\text{aaab})^2 X_1 \text{ab} \beta X_2 \subset \text{Ver}));$$

the above purpose is met, therefore, by providing that "$(\alpha^2)\mathbf{Z}$" be true if and only if, for all classes $X_1$ and $X_2$, the result of putting "$\beta \ni ((\text{aaab})^2 X_1 \text{ab}\beta X_2 \subset \text{Ver})$" for "$\alpha^2$" in "$\mathbf{Z}$" is true. This is provided by treating "$\alpha^2$" in "$\mathbf{Z}$" as an abbreviation of "$\beta \ni ((\text{aaab})^2\alpha_1\text{ab}\beta\alpha_2 \subset \text{Ver})$" and then explaining "$(\alpha)^2\mathbf{Z}$" as "$(\alpha_1)\ (\alpha_2)\mathbf{Z}$."

8.1 is next carried over to classes of classes, in this form:

$$(\mathbf{X} \in \alpha^2 \ni \mathbf{Z}) =_{\text{df}} (\alpha^2)\ ((\mathbf{X} = \alpha^2) \supset \mathbf{Z}).$$

We thus obtain symbols "$\alpha^2 \ni \mathbf{Z}$" purporting to denote classes of classes of classes. 8.4–8.6 then generate the class algebra in the third type. The immunity of the present scheme to confusion of type is exemplified by the fact, readily verified, that "$\mathbf{X} \in \alpha^2 \ni \mathbf{Z}$" does not reduce through our definitions to primitive notation unless "$\mathbf{X}$" is an incomplete symbol of the form "$\beta \ni \mathbf{W}$."

Since "$\beta^2$" means "$\gamma' \ni ((\text{aaab})^2\beta_1\text{ab}\gamma\beta_2 \subset \text{Ver})$," an ostensible sentence form "$\mathbf{W}$" in "$\beta^2$" is actually an abbreviation of a sentence form in "$\beta_1$" and "$\beta_2$." By the principle of regimentation, therefore, there will be classes $X_1$, $X_2$ and $X_3$ such that

$$(\beta^2 \ni \mathbf{W} = \beta^2 \ni ((\text{aaab})^4 X_1\text{ab}\beta_1 X_2\text{ab}\beta_2 X_3 \subset \text{Ver})).$$

Hence, analogously to the introduction of the quantification "$(\alpha^2)\mathbf{Z}$," we can introduce quantification with respect to classes of third type by explaining "$(\alpha^3)\mathbf{Z}$" as "$(\alpha_1)\ (\alpha_2)\ (\alpha_3)\mathbf{Z}$" where the "$\alpha^3$" in "$\mathbf{Z}$" is explained as

"$\beta^2 \ni ((\text{aaab})^4\alpha_1\text{ab}\beta_1\alpha_2\text{ab}\beta_2\alpha_3 \subset \text{Ver})$."

Thereupon 8.1 yields symbols "$\alpha^3 \ni \mathbf{Z}$" purporting to denote classes of fourth type; and so the process continues. The definitions introducing the higher-type classes generally are epitomized in a recursive scheme embodying 8.1, 8.5 and the following:

$$\alpha^1 =_{\text{df}} \alpha_1. \qquad 9.1. \qquad (\alpha^{n+1}) =_{\text{df}} (\alpha^n)\ (\alpha_{n+1}). \qquad 9.2.$$

$$\alpha^{n+1} =_{\text{df}} \beta^n \ni ((\text{aaab})^{2n}\alpha_1\text{ab}\beta_1\alpha_2\text{ab}\beta_2 \ldots \alpha_{n+1} \subset \text{Ver}). \qquad 9.3.$$

The theory of classes is thus at hand for all types. Relations can now be introduced on the basis of classes by Wiener's method.[9]

[1] See Quine, "Toward a Calculus of Concepts," *Jour. Symb. Logic*, **1** (1936), next to last paragraph.

[2] See Carnap, *Monatsh. für Math. u. Phys.*, **41**, 263–266 (1934).

[3] See Lukasiewicz, *C. R. Soc. Sci. et Let. Varsovie*, **24**, Cl. III, 154–156 (1931).

[4] For treatment of certain problems attending the notion of an expression thing see Tarski, *Studia Phil.*, **1**, 290–291 (1935).

[5] See C. S. Peirce, *Collected Papers*, **2**, 143.

[6] See Quine, *Mind*, N. S., **43**, 472–476 (1934).

[7] This is true for the same reason for which it is true in Lukasiewicz's scheme that no formula can be initial to another. The latter fact follows immediately from Lukasiewicz, op. cit., 156n.

[8] See Sheffer, *Trans. Amer. Math. Soc.*, **14**, 481–488 (1913).

[9] See Wiener, *Proc. Camb. Phil. Soc.*, **17**, 387–390 (1914).